Functional Analysis of a Novel Genome-Wide Association Study Signal in SMAD3 That Confers Protection From Coronary Artery Disease

Turner, Adam W.; Martinuk, Amy; Silva, Anada; Lau, Paulina; Nikpay, Majid; Eriksson, Per; Folkersen, Lasse; Perisic, Ljubica; Hedin, Ulf; Soubeyrand, Sebastien

Total number of authors:
11

Published in:
Arteriosclerosis, Thrombosis, and Vascular Biology

Link to article, DOI:
10.1161/ATVBAHA.116.307294

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Functional Analysis of a Novel Genome-Wide Association Study Signal in *SMAD3* That Confers Protection From Coronary Artery Disease

Adam W. Turner, Amy Martinuk, Anada Silva, Paulina Lau, Majid Nikpay, Per Eriksson, Lasse Folkersen, Ljubica Perisic, Ulf Hedin, Sebastien Soubeyrand, Ruth McPherson

Objective—A recent genome-wide association study meta-analysis identified an intronic single nucleotide polymorphism in *SMAD3*, rs56062135C>T, the minor allele (T) which associates with protection from coronary artery disease. Relevant to atherosclerosis, SMAD3 is a key contributor to transforming growth factor-β pathway signaling. Here, we seek to identify ≥1 causal coronary artery disease–associated single nucleotide polymorphisms at the *SMAD3* locus and characterize mechanisms whereby the risk allele(s) contribute to coronary artery disease risk.

Approach and Results—By genetic and epigenetic fine mapping, we identified a candidate causal single nucleotide polymorphism rs17293632C>T (D′, 0.97; r^2, 0.94 with rs56062135) in intron 1 of *SMAD3* with predicted functional effects. We show that the sequence encompassing rs17293632 acts as a strong enhancer in human arterial smooth muscle cells. The common allele (C) preserves an activator protein (AP)-1 site and enhancer function, whereas the protective (T) allele disrupts the AP-1 site and significantly reduces enhancer activity (P<0.001). Pharmacological inhibition of AP-1 activity upstream demonstrates that this allele-specific enhancer effect is AP-1 dependent (P<0.001). Chromatin immunoprecipitation experiments reveal binding of several AP-1 component proteins with preferential binding to the (C) allele. We show that rs17293632 is an expression quantitative trait locus for *SMAD3* in blood and atherosclerotic plaque with reduced expression of *SMAD3* in carriers of the protective allele. Finally, siRNA knockdown of *SMAD3* in human arterial smooth muscle cells increases cell viability, consistent with an antiproliferative role.

Conclusions—The coronary artery disease–associated rs17293632C>T single nucleotide polymorphism represents a novel functional cis-acting element at the SMAD3 locus. The protective (T) allele of rs17293632 disrupts a consensus AP-1 binding site in a SMAD3 intron 1 enhancer, reduces enhancer activity and SMAD3 expression, altering human arterial smooth muscle cell proliferation. **(*Arterioscler Thromb Vasc Biol*. 2016;36:972-983. DOI: 10.1161/ATVBAHA.116.307294.)**

Key Words: binding sites ■ coronary artery disease ■ genome-wide association study ■ genomics ■ SMAD3 protein

Coronary artery disease (CAD) characterized by the deposition of atherosclerotic plaque in the epicardial arteries remains a leading cause of death and disability.[1] Environmental variables contribute to CAD but genetic factors are thought to be of equal importance. During the past several years, large consortia have conducted meta-analyses of genome-wide association studies (GWAS) for CAD[2,3] resulting in the identification of >45 CAD-associated loci.[3] Together, these earlier studies explained ≈10.6% of the predicted heritability of CAD and for the majority, the functional link to atherosclerosis remains unknown.

Recently, using phased haplotypes from the 1000 Genomes Project to impute an enlarged number of single nucleotide polymorphisms (SNPs) for 48 GWAS, consisting of >180 000 cases and controls, we identified an additional 10 CAD-associated loci. One of the novel CAD-associated SNPs is rs56062135C>T at the *SMAD3* locus on chromosome 15, and the common allele (C) is associated with increased risk (odds ratio=1.07) relative to the (T) allele (Table 1).[4] Of note, this SNP is not in linkage disequilibrium (LD) with rs17228212T>C, reported to be associated with CAD in an earlier GWAS.[5]

The SMAD3 transcription factor is an important mediator of transforming growth factor-β (TGF-β) signaling, regulating transcription of genes with SMAD-binding elements. On TGF-β stimulation, the type II TGF-β receptor recruits and phosphorylates the type I TGF-β receptor,

Received on: January 28, 2016; final version accepted on: February 19, 2016.

From the Atherogenomics Laboratory (A.W.T., A.M., A.S., P.L., S.S., R.M.) and Department of Medicine, Ruddy Canadian Cardiovascular Genetics Centre (M.N., R.M.), University of Ottawa Heart Institute, Ottawa, Canada; Atherosclerosis Research Unit (P.E., L.F.) and Department of Molecular Medicine and Surgery (L.P., U.H.), Karolinska University Hospital, Stockholm, Sweden; and Department of Systems Biology, Technical University of Denmark, Copenhagen, Denmark (L.F.).

This manuscript was sent to Anne Tybjærg-Hansen, Consulting Editor, for review by expert referees, editorial decision, and final disposition.

The online-only Data Supplement is available with this article at http://atvb.ahajournals.org/lookup/suppl/doi:10.1161/ATVBAHA.116.307294/-/DC1.

Correspondence to Ruth McPherson, MD, PhD, Atherogenomics Laboratory, University of Ottawa Heart Institute, 40 Ruskin St, H4203, Ottawa, Canada K1Y 4W7. E-mail rmcpherson@ottawaheart.ca

Nonstandard Abbreviations and Acronyms

AP-1	activator protein-1
CAD	coronary artery disease
ChIP	chromatin immunoprecipitation
eQTL	expression quantitative trait locus
GWAS	genome-wide association study
hASMCs	human arterial smooth muscle cells
LD	linkage disequilibrium
qRT-PCR	quantitative reverse transcription polymerase chain reaction
SMCs	smooth muscle cells
SNP	single nucleotide polymorphism
TGF-β	transforming growth factor-β

which subsequently phosphorylates the SMAD3 protein. Phosphorylated SMAD3 then forms a complex with the common SMAD4 that subsequently translocates to the nucleus and regulates transcription.[6,7] Relevant to a role in atherosclerosis, in systems genetics analysis of multiple GWAS, we identified TGF-β signaling and SMAD transcriptional activities as enriched pathways for CAD association.[8] However, despite extensive data on the functions of TGF-β with respect to atherosclerosis,[9,10] the roles of SMAD proteins particularly SMAD3 and SMAD3 signaling are less well-understood.

SMAD3 is expressed at low levels in healthy human aorta by immunohistochemistry and quantitative reverse transcription polymerase chain reaction (PCR).[11] There is, however, a marked increase in SMAD3 and other SMAD proteins in fibrofatty lesions, with expression mostly limited to CD68-positive macrophages/macrophage-derived foam cells in these samples. Conversely, in fibrous atherosclerotic plaques, there are high levels of SMAD3 in vascular smooth muscle cells (SMCs), suggesting that the role of SMAD3 in atherosclerosis depends on cell type and stage of atherosclerosis.[11] Higher SMAD3 expression in SMCs of the fibrous plaque coincides with TGF-β–mediated synthesis of collagen and other extracellular matrix proteins, which contribute to plaque stability.[12] Rare *SMAD*3 mutations cause aneurysms–osteoarthritis syndrome, an autosomal dominant form of thoracic aortic aneurysms and dissections.[13]

Here, we have identified and characterized a functional SNP, rs17293632C>T, in *SMAD3* that resides in a newly identified GWAS locus for CAD. Consistent with functionality, we show that rs17293632C>T is an expression quantitative trait locus (eQTL) in both whole blood and carotid plaque samples with reduced expression of *SMAD3* in carriers of the protective (T) allele. We demonstrate that the protective (T) allele disrupts a consensus activator protein (AP)-1 binding site within a strong intronic enhancer and impairs enhancer activity in human arterial SMCs (hASMCs), as well as HeLa and HepG2 cells. Pharmacological blocking of AP-1 signaling and of DNA binding ability markedly attenuates enhancer activity, and chromatin immunoprecipitation (ChIP) experiments reveal binding of numerous AP-1 components to the region encompassing rs17293632, with impaired binding to the (T) versus (C) allele. Finally, siRNA knockdown of *SMAD3* in hASMCs increases cell viability, consistent with an antiproliferative role.

Materials and Methods

Materials and Methods are available in the online-only Data Supplement.

Results

Haplotype Analysis of CAD-Associated SNPs at the *SMAD3* Locus and Identification of Candidate Functional SNPs

The index GWAS SNP rs56062135C>T is intronic to *SMAD3* on chromosome 15 (Table 1).[4] The *SMAD3* gene is ≈130 kb in size with 9 exons, mostly clustered toward the 3′ end of the gene (Figure 1A). The regional association plot of *SMAD3* and chromosome 15 indicated a cluster of CAD-associated SNPs in *SMAD3* in proximity to rs56062135C>T, with weaker CAD-associated signals in nearby genes at this locus (Figure 1B). Our conditional and joint association analysis showed that the association between *SMAD3* and CAD is mainly explained by rs56062135 because it was the only SNP that remained significant in the final joint model (Table III in the online-only Data Supplement). Furthermore, we investigated the effect of rs56062135 on association of *SMAD3* variants with CAD by doing conditional analysis, as presented in Table V in the online-only Data Supplement. None of the variants tagging *SMAD3* reach GWAS significance ($P<5\times10^{-8}$) after conditional analysis, and the vicinity SNPs to rs56062135 showed GWAS significant association in the CARDIoGRAM meta-analysis because they are in high LD with rs56062135 ($r^2>0.9$).

We then sought to identify other SNPs in strong LD with this SNP using HaploReg v3 (http://www.broadinstitute.org/mammals/haploreg/haploreg_v3.php) and Haploview (https://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview).[14] This analysis revealed 6 other intronic SNPs in

Table 1. Association of the Genome-Wide Association Study–Reported rs56062135[4] SNP and the rs17293632 SNP in Intron 1 of *SMAD3* With Coronary Artery Disease

Lead Variant	Locus Name	Chr	A1/A2	Effect Allele (A1) Frequency	Imputation Quality	I^2	Heterogeneity *P*	N Studies	Association Model			
									Additive		Recessive	
									OR (95% CI)	*P*	OR (95% CI)	*P*
rs56062135	SMAD3	15	C/T	0.79	0.98	0.00	0.67	46	1.07 (1.05–1.10)	4.50E–09	1.17 (1.10–1.25)	8.88E–07
rs17293632	SMAD3	15	C/T	0.79	0.98	0.00	0.76	46	1.07 (1.05–1.10)	5.72E–09	1.17 (1.11–1.23)	1.22E–06

Chr indicates chromosome; CI, confidence interval; OR, odds ratio; and SNP, single nucleotide polymorphism.

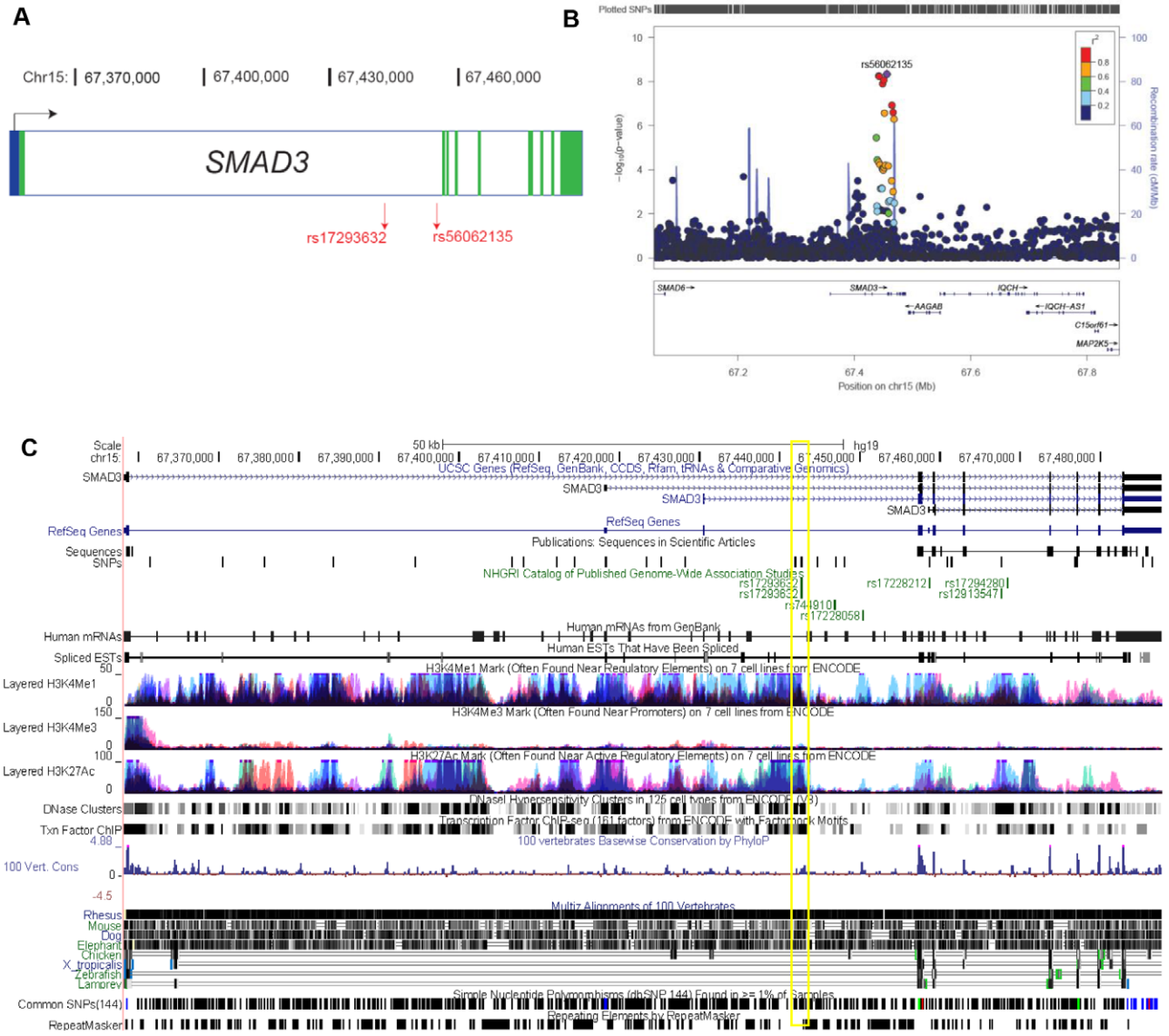


Figure 1. Overview and chromatin signatures of the *SMAD3* locus on chromosome 15. **A**, *SMAD*3 is found on the long arm of chromosome 15 and contains 9 exons. Single nucleotide polymorphisms (SNPs) denoted in red represent the index rs56062135 GWAS SNP and the functional rs17293632 candidate SNP. **B**, Regional association plot for coronary artery disease (CAD) association for *SMAD3* and proximal genes on chromosome 15 generated from 1000 genomes project imputation data. **C**, University of California Santa Cruz (UCSC) genome browser annotation and ENCODE project data for the entire *SMAD3* gene on chromosome 15. SNPs from previously published genome-wide association studies are denoted in green. Chromatin regulatory features such as DNase hypersensitivity sites, histone modifications (H3K4me1, H3K4me3, H3K27ac), and predicted transcription factor binding from chromatin immunoprecipitation–sequencing data are all included. EST indicates expressed sequence tag; and NHGRI, National Human Genome Research Institute.

strong LD (r^2>0.8), including rs17293632C>T within intron 1 (Table 2). No *SMAD3* exonic SNPs or SNPs predicted to alter splicing were in LD with rs56062135. Haploview software and imported linkage data from the 1000 Genomes Browser reveals 20 separate LD blocks at the *SMAD3* locus; rs56062135C>T and linked SNPs are all found within block 13 (Figure I in the online-only Data Supplement). This fine-mapping effort led us to conclude that the functional SNP(s) at the *SMAD3* locus are noncoding.

We next attempted to narrow down candidate functional/causal SNPs at this locus. We conducted epigenetic fine-mapping and scanned publicly available databases to determine whether any of these intronic SNPs lie within a noncoding gene regulatory element such as an enhancer or promoter.[15–17] To help localize truly functional variants among a large pool of SNPs, we looked up rs56062135C>T and linked SNPs in the RegulomeDB (http://www.regulomedb.org/index) database. RegulomeDB contains experimental data sets from ENCODE and other sources in addition to computational predictions of regulatory potential.[18] RegulomeDB guides interpretation of regulatory variants and assigns each variant a score, 1 being likely to be functional and 6 unlikely to be functional. This analysis revealed rs17293632C>T as the only *SMAD3* SNP to represent a suitable candidate functional/causal variant

Table 2. SNPs in Intron 1 of *SMAD3* in Strong Linkage Disequilibrium (r^2>0.8) With rs56062135 and RegulomeDB Scores for Predicted Function

SNP	Chromosome	Position (hg19)	D′	r^2	RegulomeDB Score
rs72743461	15	67441750	0.97	0.92	4
rs17293632	15	67442596	0.97	0.94	2a
rs56375023	15	67448363	0.98	0.95	…
rs17228058	15	67450305	0.99	0.98	5
rs56062135	15	67455630	1.00	1.00	…
rs72743477	15	67464291	0.98	0.93	5
rs72743482	15	67466599	0.97	0.91	5

RegulomeDB scores for predicted function (1=likely to be functional; and 6=not likely to be functional). Linkage disequilibrium values were obtained from the HaploReg (v3). SNP indicates single nucleotide polymorphism.

(Table 2). ENCODE and regulatory analysis demonstrated that rs17293632C>T is located in a gene regulatory hotspot (Figure 1C). HaploReg data indicates rs17293632C>T is located in a region with enhancer histone marks in 8 cell types, DNase hypersensitivity sites in 41 cells types, and alters 25 in silico transcription factor–binding sites. ChIP-Seq data from ENCODE reveal binding of dozens of transcription factors to the DNA region encompassing rs17293632C>T (Figure II in the online-only Data Supplement). The rare allele (T; MAF=0.21) of rs17293632C>T is protective (P=5.72×10^{-9}; Table 1), comparable with the rs56062135C>T GWAS index SNP. In contrast, rs56062135 and rs56375023 had no available RegulomeDB data (Table 2), and ENCODE data did not suggest causal effects.

rs17293632 (T) Disrupts the Function of a Novel Enhancer in Intron 1 of *SMAD3*

We next sought to functionally verify the presence of an enhancer by investigating the region containing rs17293632C>T and 1 kb 5′ and 3′ flanking sequences (Figure III in the online-only Data Supplement). This 2-kb sequence contains several SNPs in addition to rs17293632C>T, as shown in Figure 2A. Pairwise LD values between SNPs in this region are shown in Figure IV in the online-only Data Supplement. We cloned this 2 kb sequence for dual luciferase assays by PCR amplification from genomic DNA homozygous CC at rs17293632 and homozygous for all other SNPs. This sequence was inserted into the pGL3-Promoter vector (Promega) that contains an SV40 promoter upstream of the firefly luciferase gene and is used to evaluate putative enhancer sequences (Figure 2B). Because enhancer effects are commonly cell type–specific,[19,20] experiments were performed in primary hASMCs with relevance to CAD. When the 2 kb sequence from *SMAD3* intron 1 was cloned into pGL3-Promoter with the C allele present for rs17293632, we observed a 10-fold increase in luciferase expression relative to no insert (P<0.001; Figure 2C). When the rs17293632 allele (C) was changed to (T), via site-directed mutagenesis, the activity of this *SMAD3* enhancer was reduced by half (P<0.001). This indicated that in hASMCs, rs17293632(T) disrupts binding of ≥1 transcription factors or chromatin-modifying enzymes. In HeLa and HepG2 cells, we found this 2-kb sequence to also act as a strong enhancer to moderate enhancer with an ≈4-fold increase in luciferase activity compared with empty vector (P<0.001; Figure 2C). Interestingly, the effects of this enhancer in HeLa and HepG2 cells are almost completely dependent on rs17293632C>T genotype because the presence of the (T) allele nearly fully disrupts enhancer activity. To further examine the contribution of the enhancer in the *SMAD3* context, we cloned a 347-bp subset of this enhancer (Figure 2A) with both alleles of rs17293632 into the SMAD3p-Luc plasmid containing the human *SMAD3* promoter in the pGL3-Basic backbone[21] (Figure 2D). The enhancer was inserted 3′ of both the *SMAD3* promoter sequence and firefly luciferase gene to mimic the orientation in the genome. As seen with the longer enhancer above, this 347-bp enhancer acts as an enhancer in the context of the *SMAD3* promoter in hASMCs, and the rs17293632 (T) allele almost completely disrupts enhancer function (P<0.001; Figure 2E).

Next, we sought to verify that the *SMAD3* intron 1 sequence containing rs17293632C>T acts as an active enhancer in hASMCs by performing ChIP with an H3K27ac antibody. Relative to nonspecific IgG, H3K27ac was highly enriched at rs17293632 (P<0.05) and suggests that this enhancer is functionally active in these cells (Figure 2F). The histone mark H3K27ac is typically deposited by the protein P300, a transcriptional coactivator commonly found at enhancers,[22,23] that ENCODE predicted to strongly bind rs17293632 in other cell types. Indeed, binding of P300 (P<0.01) to the rs17293632 sequence was enriched in the hASMC system (Figure 2F), further supporting the presence of a functional enhancer.

SMAD3 rs17293632C>T Is an eQTL in Whole Blood and Human Atherosclerotic Plaque Tissue

An eQTL effect provides strong evidence for a functional effect of a given SNP.[24,25] We first searched common publicly available online eQTL databases to determine whether rs17293632C>T or LD SNPs are associated with transcript levels of *SMAD3* or another gene. We originally queried both the eQTL Browser (http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/) from the University of Chicago and the NCBI eQTL Browser (http://www.ncbi.nlm.nih.gov/projects/gap/eqtl/index.cgi), yet neither rs17293632C>T nor its linked SNPs appeared as eQTLs. Furthermore, we could not identify any reported eQTL SNP within the *SMAD3* gene in an electronic search. We then sought to determine whether rs17293632C>T

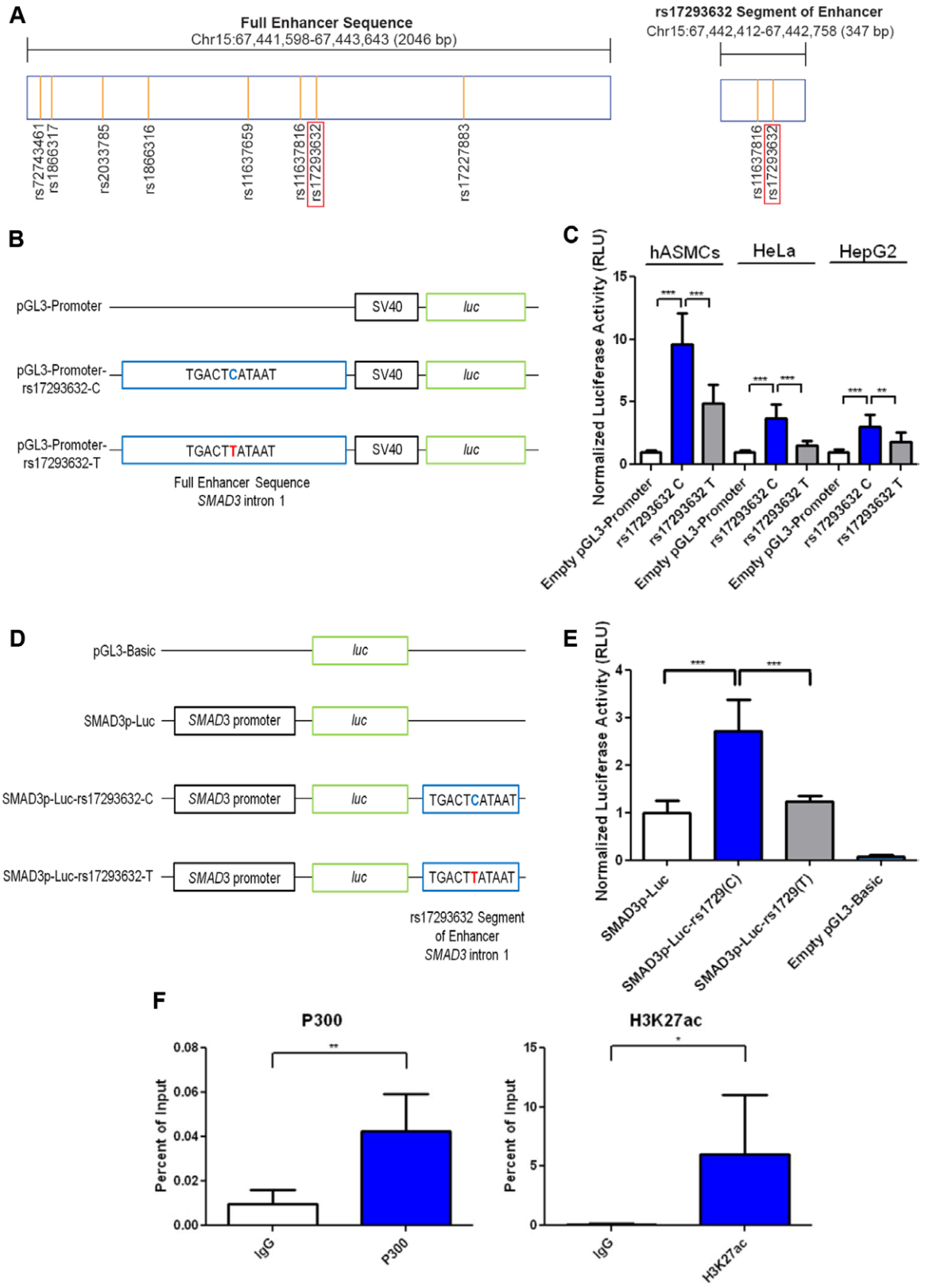


Figure 2. rs17293632 resides within a strong enhancer in human arterial smooth muscle cells (hASMCs) and other cell types and is responsible for transcriptional activity of this enhancer. **A**, Schematic of the different enhancer sequences containing rs17293632 investigated in dual luciferase assays. Single nucleotide polymorphisms located within these enhancer sequences are denoted by their relative position. **B**, Overview of the *SMAD3* enhancer reporter constructs with the SV40 promoter used for dual luciferase assays. These constructs contain the putative ≈2 kb full *SMAD3* enhancer sequence (containing either the [C] or [T] allele at rs17293632) upstream of the SV40 promoter and firefly luciferase gene of the pGL3-Promoter vector that allows for identification of functional enhancer sequences. Empty pGL3-Promoter vector was used as a negative control. **C**, Firefly luciferase values for the SV40 promoter *SMAD3* enhancer (*Continued*)

Figure 2 Continued. reporter constructs transfected in HITC6 hASMCs, HeLa cells, and HepG2 cells. Firefly luciferase values were normalized to cotransfected *Renilla* luciferase. The hASMCs were harvested for dual luciferase assays 48 hours after transfection and HeLa and HepG2 cells harvested 24 hours after transfection. Plotted values represent mean±SD of 3 independent experiments (triplicates each experiment); ***P<0.001; **P<0.01; ANOVA with Bonferroni post hoc test. **D**, Overview of the *SMAD3* enhancer reporter constructs with the *SMAD3* promoter used for dual luciferase assays. Enhancer segments, 347 bp, (rs17293632 allele either [C] or [T]) were inserted downstream of the *SMAD3* promoter and firefly luciferase gene in the SMAD3p-Luc plasmid. **E**, Firefly luciferase values for the *SMAD3* promoter *SMAD3* enhancer reporter constructs that were transfected in HITC6 hASMCs and harvested 48 hours after transfection, with raw firefly luciferase values normalized using cotransfected *Renilla* luciferase. Plotted values represent mean±SD of 3 independent experiments (triplicates each experiment); ***P<0.001; ANOVA with Bonferroni post hoc test. **F**, Chromatin immunoprecipitation data from HITC6 hASMCs homozygous CC at rs17293632 shows enrichment of both P300 protein as well as the H3K27ac chromatin mark at the rs17293632 intronic sequence compared with nonspecific IgG. Values are from 3 independent experiments and displayed as percentage of input DNA. Error bars denote SD; **P<0.01, *P<0.05, 2- tailed unpaired Student t test.

acts as an eQTL for *SMAD3* expression using PAXgene-derived mRNA from whole blood of healthy individuals. Imputed rs17293632C>T genotypes from the Affymetrix 6.0 array were verified experimentally using a TaqMan SNP Genotyping Assay specific for rs17293632. Because a previous study reported sex-specific effects on *SMAD3* expression,[26] we selected an even number of male and female subjects across the 3 genotypes for rs17293632C>T and all were matched for age and body mass index. We detected a significant association between rs17293632 genotype and *SMAD3* mRNA levels in whole blood (P<0.05; 1-way ANOVA; Figure 3A).

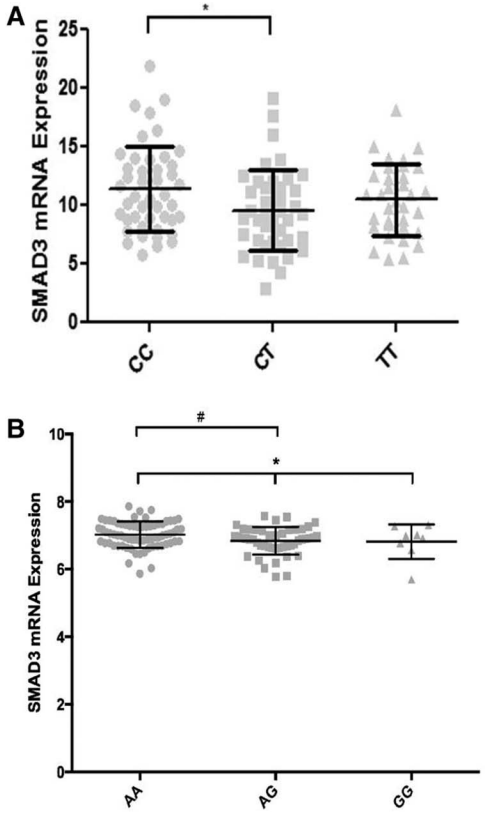


Figure 3. rs17293632 represents a *cis*-expression quantitative trait locus for *SMAD3* expression in human whole blood and carotid plaque tissue. **A**, *SMAD3* mRNA levels in whole blood was measured across the 3 genotypes for rs17293632 by quantitative reverse transcription polymerase chain reaction and normalized to expression levels of *PPIA*. The number of patients for the CC, CT, and TT genotypes were 50, 50, and 41, respectively. Whole blood RNA was extracted from PAXgene Blood RNA tubes taken from fasting healthy individuals matched for age and sex. Lines on the graph represent mean±SD; *P<0.05; ANOVA with Bonferroni post hoc test. **B**, Association of rs16950687, proxy single nucleotide polymorphism (SNP) for rs17293632, with *SMAD3* mRNA expression in n=125 genotyped carotid plaque samples from the Biobank of Karolinska Endarterectomies (BiKE) study; *P<0.05; 1-way ANOVA; #P<0.05; t test between the AA and AG genotypes.

We next queried whether rs17293632C>T associates with *SMAD3* expression in human atherosclerotic carotid plaque samples.[27–30] Microarray data from plaque tissue in the Biobank of Karolinska Endarterectomies (BiKE) cohort, available for n=125 patients, were analyzed for association with rs17293632C>T. Because neither this SNP nor rs56062135C>T were present on the chip, rs16950687A>G (D′ 0.949; r^2 0.859 with rs17293632; D′ 1; r^2 0.953 with rs56062135) was used as a proxy SNP. As shown, rs16950687A>G was significantly associated with *SMAD3* mRNA levels in carotid lesions (P<0.05; Figure 3B). We also examined association of rs17293632C>T with *SMAD3* expression in several different tissues from the Advanced Study of Aortic Pathology.[31] eQTL data from the ascending aorta adventitia were directionally consistent with the effects in blood and plaque tissue, but did not reach significance. Taken together, these data provide support for rs17293632C>T as an eQTL for *SMAD3* mRNA expression in human whole blood and plaque tissue, consistent with an effect on gene regulation.

Enhancer Activity at Intron 1 of *SMAD3* Is AP-1 Dependent

Although numerous transcription factors were predicted to bind the *SMAD3* intron 1 enhancer, the protective allele (T) of rs17293632 is predicted to disrupt the consensus binding sequence, 5′-TGA[G/C]TCA-3′, for the transcription factor AP-1 (Figure 4A). To investigate the role of AP-1 at the *SMAD3* intron 1 enhancer, we first sought to determine the effects of inhibiting AP-1 signaling pharmacologically using SP600125, an inhibitor of the Jun amino-terminal kinases (JNK) pathway that blocks both the phosphorylation and transcriptional activity of c-Jun, a key player in AP-1–mediated transcription.[32] In hASMCs, blocking AP-1 activation via the JNK pathway was shown to strongly diminish enhancer activity of the full *SMAD3* intron 1 reporter containing the protective (C) allele at rs17293632 (Figure 4B). As predicted, SP600125 did not alter the enhancer activity of the construct containing rs17293632 (T), where the AP-1 site is already disrupted. Reduced c-Jun phosphorylation in response to SP600125 treatment was confirmed via Western blot (Figure V in the online-only Data Supplement). These data demonstrate that the AP-1 transcription factor is crucial for proper enhancer function at this specific location.

Next, we sought to determine whether activating AP-1 increases endogenous SMAD3 mRNA levels by using the

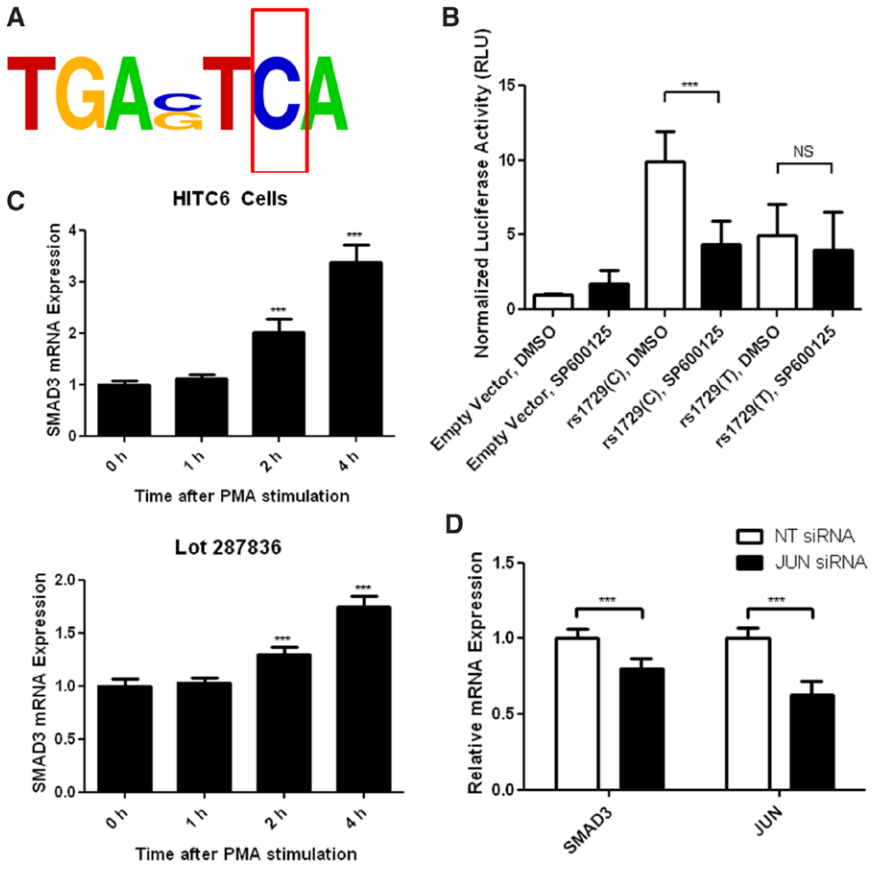


Figure 4. Activator protein (AP)-1 regulates enhancer activity at *SMAD3* intron 1 and expression of SMAD3 mRNA. **A**, Depiction of the consensus AP-1 binding site, 5′-TGA[G/C]TCA-3′, with the location of the rs17293632 single nucleotide polymorphism highlighted in the red box. **B**, HITC6 human arterial smooth muscle cells (hASMCs) were pretreated with 50 μmol/L SP600125, which blocks phosphorylation of c-Jun, for 1 hour before transfection with the full-length *SMAD3* enhancer reporter vectors or empty pGL3-Promoter vector for 24 hours. Values are plotted as mean±SD of 3 independent experiments (triplicates each experiment); NS indicates not significant; ***P<0.001; ANOVA with Bonferroni post hoc test. **C**, HITC6 and lot 287836 hASMCs were treated with a final concentration of 50 nmol/L phorbol 12-myristate 13-acetate (PMA) and harvested for RNA at 0, 1, 2, and 4 hours. *SMAD3* mRNA levels were assessed by quantitative reverse transcription polymerase chain reaction (qRT-PCR) and normalized to expression levels of β-actin. Values represent mean±SD of 3 independent experiments (triplicates each experiment); ***P<0.001 compared with 0 hours; ANOVA with Bonferroni post hoc test. **D**, HITC6 hASMCs were transfected with either 50 nmol/L nontarget siRNA or 50 nmol/L *JUN* siRNA for 48 hours. *SMAD3* and *JUN* mRNA levels were assessed by qRT-PCR and normalized to expression levels of β-actin. Values represent mean±SD of 3 independent experiments (triplicates each experiment); ***P<0.001 compared with nontarget (NT) siRNA; 2- tailed unpaired Student *t* test.

commonly used phorbol ester PMA (phorbol 12-myristate 13-acetate, also referred to as 12-*O*-Tetradecanoylphorbol-13-acetate/TPA). During the course of 1 to 4 hours PMA stimulation, SMAD3 mRNA levels increase in 2 separate lots of hASMCs (Figure 4C). Finally, knocking down c-Jun expression via siRNA decreases SMAD3 mRNA during the course of 48 hours (P<0.001), suggesting that c-Jun and AP-1 normally increase SMAD3 transcription (Figure 4D). In contrast, knocking down other AP-1 components such as c-Fos, JunB, and JunD did not produce significant changes in SMAD3 mRNA (data not shown).

Allele-Specific Binding of AP-1 Components for rs17293632 in Primary Arterial SMCs

ENCODE ChIP-Seq data (Figure II in the online-only Data Supplement) indicate that AP-1 proteins c-Fos, FosL1, c-Jun, JunB, and JunD are able to bind the rs17293632 site in the *SMAD3* intron but with little data on vascular cell types. We thus performed ChIP experiments in HITC6 hASMCs homozygous for the C allele. Protein-DNA complexes were pulled down using antibodies specific for c-Fos, c-Jun, JunB, and JunD, as well as a nonspecific IgG control. We demonstrated strong binding of c-Fos and JunD to rs17293632(C), as well as moderate binding of JunB and c-Jun (Figure 5A). AP-1 ChIP products were also amplified for negative control intronic genomic regions (Figure VI in the online-only Data Supplement). Relative to the input DNA, we observed much weaker binding of AP-1 proteins to the negative control genomic areas (intronic sequences selected that are not predicted to bind AP-1).

We next sought to determine whether changing rs17293632 (C) to (T) affects binding of AP-1 proteins by performing allele-specific ChIP-quantitative PCR, using lot 287836 primary hASMCs that were heterozygous for rs17293632. To assess the allele-specific binding of AP-1, we first performed ChIP and followed with a nested PCR-TaqMan genotyping method for quantification. Protein/DNA complexes were immunoprecipitated with c-Fos, c-Jun, JunB, and JunD antibodies. After PCR amplification of the eluted ChIP DNA,

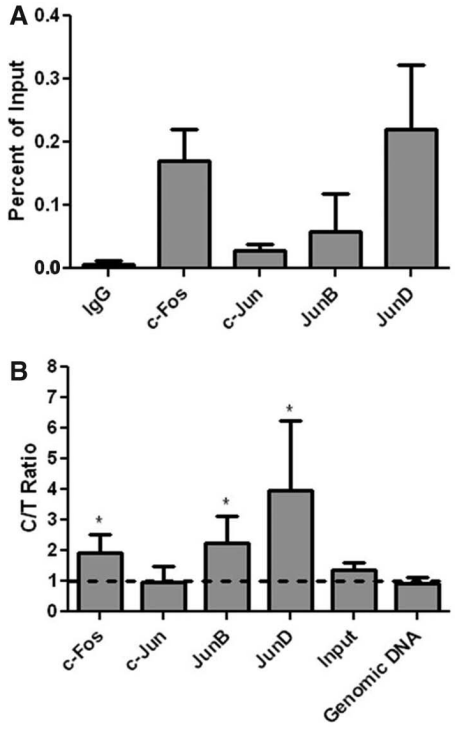


Figure 5. Activator protein (AP)-1 proteins bind to the rs17293632 sequence and preferentially bind the common (C) allele. **A**, Chromatin immunoprecipitation data from HITC6 human arterial smooth muscle cells (hASMCs) show enrichment of AP-1 component proteins c-Fos, c-Jun, JunB, and JunD at the intronic enhancer comprising rs17293632. Values on the *y* axis are expressed as percentage of the input chromatin DNA and are representative of 3 independent experiments. Error bars denote SD. **B**, Allele-specific chromatin immunoprecipitation (ChIP) was conducted in lot 287836 hASMCs, heterozygous at rs17293632. Eluted ChIP DNA values are expressed as the ratio of the rs17293632 (C) to (T) allele and are representative of 3 independent ChIP/Polymerase Chain Reaction/TaqMan experiments. *$P<0.05$; *t* test compared with genomic DNA TaqMan control. Error bars denote SD.

TaqMan genotyping was performed to determine the ratios of the (C) to (T) alleles for rs17293632 in the amplified ChIP products. These TaqMan values were then applied to a standard curve consisting of different ratios of the 2 alleles (generated from homozygous genotypes). We found strong enrichment of (C) allele binding for both c-Fos and JunD ($P<0.05$; Figure 5B), both AP-1 members bound the strongest to this region as shown in Figure 5A. JunB also demonstrated enrichment for the C allele ($P<0.05$), whereas c-Jun showed, surprisingly, no enrichment (as described in Discussion). As expected, hASMC lot 287836 input and genomic DNA control samples had C:T ratios ≈1.0. Overall, these allele-specific ChIP data demonstrate enriched binding of AP-1 to the common (C) versus protective (T) allele, consistent with the above luciferase and eQTL data. Altered binding of AP-1 would lower the efficacy of the *SMAD3* intron 1 enhancer and thus reduce *SMAD3* transcription and mRNA levels.

SMAD3 Knockdown Increases hASMC Proliferation

SMAD3 knockdown in normal SMC medium resulted in a 15% to 25% increase in viable cells ($P<0.01$; Figure 6) with similar results for both lots of hASMCs. These findings indicate that SMAD3 has a negative effect on viability in hASMCs, which likely reflects an effect on cell proliferation. Because previous studies have shown that TGF-β induces apoptosis via a SMAD-dependent mechanism,[33,34] we determined protein levels of PARP (Poly [ADP-ribose] polymerase), a DNA repair protein, that is a key caspase target. Caspase-mediated cleavage of PARP is a common marker of apoptosis, detectable by Western blotting because of the conversion of the 116 kDa full-length protein to its 89 kDa form.[35] We could not detect the characteristic of 89 kDa form of PARP in nontarget siRNA-treated hASMCs, and observed no difference in this marker of apoptosis between nontarget and *SMAD3* knockdown samples (Figure VII in the online-only Data Supplement). Furthermore, the microscopic appearance of hASMCs did not differ between nontarget siRNA and *SMAD3* siRNA treatment, supportive of an increase in hASMC proliferation in response to *SMAD3* siRNA knockdown (Figure 6). Thus, these data indicate that SMAD3 knockdown is not repressing a putative underlying apoptosis but rather suggest that SMAD3 regulates cell proliferation via other means.

Discussion

Despite the success of recent GWAS, there has been limited progress in understanding the function of the multiple risk loci identified. Multiple hurdles exist including the fact that most are located in intergenic regions or introns rather than coding sequences. Second, the identified polymorphism is most often not causal but rather in LD with a neighboring or even distal causal polymorphism. Third, although the association is at the level of genomic DNA, the relevance may be restricted to a particular tissue or organ. Finally, the direct contribution of any given locus could be temporally restricted; for instance, early transient activation followed by long-standing enduring epigenetic repercussions.

Here, we have investigated a recently discovered GWAS signal near rs56062135C>T in *SMAD3*[4] that is not linked to a previously reported GWAS signal at rs17228212C>T.[5] By genetic and epigenetic fine mapping, we identified a SNP, rs17293632C>T, in almost perfect LD with rs56062135C>T, in a region with chromatin histone marks suggestive of enhancer activity. Functional studies confirmed that rs17293632C>T is located within a strong enhancer sequence in primary hASMCs, HeLa, and HepG2 cells. The common (C) allele maintains the conserved AP-1 binding site and results in greater AP-1 enhancer binding and maximal enhancer activity, whereas the protective (T) allele (MAF 0.21) disrupts the AP-1 recognition sequence, reduces AP-1 binding, and impairs enhancer activity. We propose that this enhancer is active in numerous cell types and that rs17293632(T) disrupts enhancer function via a universal mechanism. Stimulation with the AP-1 activator PMA increases endogenous SMAD3 mRNA levels in hASMCs, whereas c-Jun siRNA knockdown decreases SMAD3 mRNA, suggesting that AP-1 positively modulates SMAD3 transcription.

Although AP-1 regulation is complex and its transcriptional activity can be stimulated by numerous factors, we observed that the presence of 5% serum was sufficient to produce an enhancer effect at *SMAD3* intron 1 likely because of

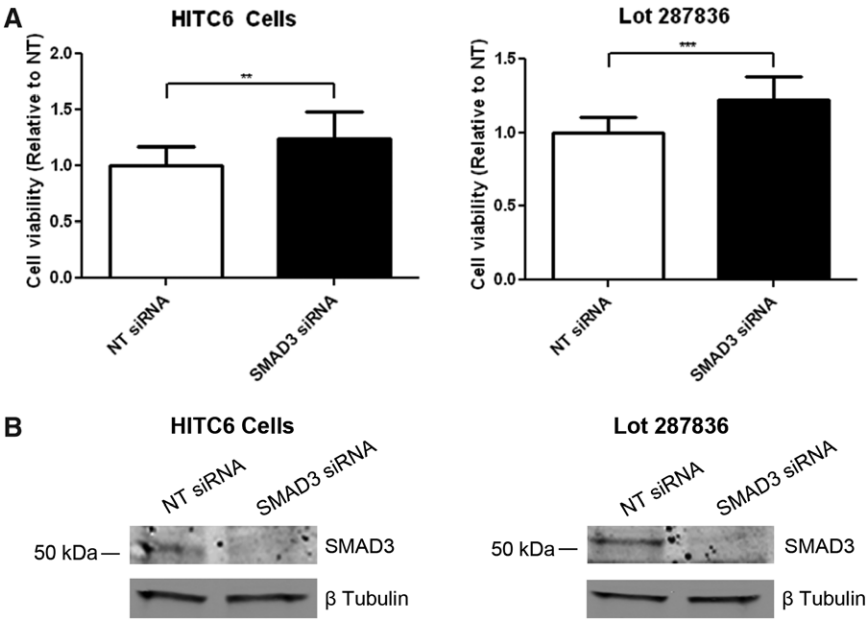


Figure 6. *SMAD3* knockdown in 2 separate lots of primary human arterial smooth muscle cells increases cell proliferation/viability. **A**, HITC6 and lot 287836 cells were transfected with either 20 nmol/L nontarget (NT) siRNA or 20 nmol/L *SMAD3* siRNA and harvested for the MTT assay 72 hours later to evaluate cell proliferation/cell viability. Values displayed represent the means±SD of 3 independent experiments; ***P<0.001; **P<0.01; ANOVA with Bonferroni post hoc test. **B**, *SMAD3* knockdown at the protein level was confirmed via Western blot, with human β-tubulin used as the loading control.

the presence of various growth factors and cytokines. ChIP data (Figure II in the online-only Data Supplement) demonstrate highly enriched binding in HITC6 cells of the AP-1 components c-Fos and JunD, with lesser but still enriched binding of c-Jun and JunB compared with nonspecific IgG. This suggests that in the presence of 5% fetal bovine serum, AP-1 is able to sufficiently bind the rs17293632(C) sequence. We were surprised not to find more enriched binding of c-Jun, a central component of all AP-1 complexes,[36,37] to the *SMAD3* intron 1 enhancer, especially because c-Jun siRNA knockdown lowers SMAD3 mRNA.

Although c-Jun is important in the transcriptional activity of AP-1, it is possible that our system did not have the proper stimulatory context to maximize c-Jun binding to DNA. Furthermore, c-Jun expression is low in many cell types and its expression requires elevation by appropriate stimuli.[36] The c-Jun antibody used may also not have been optimal for ChIP experiments. In addition, the epitope recognized by the c-Jun antibody may not be accessible in our ChIP material because of the potential of the formaldehyde cross-linking of protein to DNA masking effect. In contrast, c-Fos and JunD bound strongly to the rs17293632 sequence in an allele-dependent fashion, yet their siRNA knockdowns did not affect SMAD3 expression. Potential reasons for this are that the c-Fos and JunD proteins are less potent transcription activators compared with c-Jun or that there is functional redundancy between some AP-1 family members.[38,39]

Here, we also demonstrated ChIP-enriched binding of P300 at the rs17293632 enhancer. P300 is a transcriptional coactivator by acting as a histone acetyltransferase that is able to relax chromatin structure and facilitate binding of other transcription factors and can also act as an adaptor protein, suggesting that P300 may help the rs17293632 *SMAD3* intron 1 enhancer loop around and interact with/activate the *SMAD3* promoter area that is located in the 5′ direction. Mechanisms of P300 recruitment to this region, roles of other transcription factors from Figure 1D ChIP-Seq data, and effect of rs17293632 on chromatin modifications warrant future investigation.

Luciferase assays consistently confirmed an effect of rs17293632C>T on enhancer activity. However, our data with the 2-kb enhancer sequence indicated changing the allele from (C) to (T) reduced reporter activity significantly but not completely relative to the empty pGL3-Promoter. Similarly, treatment with SP600125 reduced the activity of the *SMAD3* reporter with the rs17293632 C allele, to the same levels as the *SMAD3* reporter with the T allele but both had greater activity in comparison with empty pGL3-Promoter, and treatment with SP600125 had no effect when the AP-1 site was disrupted. Thus, the 2-kb *SMAD3* intron 1 sequence investigated still has some enhancer activity in hASMCs independent of the rs17293632 SNP studied here. As indicated in Figure III in the online-only Data Supplement, other transcription factors are predicted to bind a few hundred base pairs away from the cluster of transcription factors that bind to the immediate rs17293632 SNP. ENCODE data also demonstrated H3K4me1 and H3K27ac enhancer histone marks across the entire 2-kb intronic sequence studied. Nevertheless, our results indicate that rs17293632 contributes to a large fraction, perhaps the majority, of this intronic enhancer regulation.

During the course of these experiments, another report highlighted rs17293632C>T as a strong candidate functional SNP. Farh et al[40] developed a fine-mapping algorithm to identify candidate casual variants for numerous autoimmune diseases and provide ChIP-Seq data for AP-1 in HeLa cells heterozygous at rs17293632. They observed robust binding of AP-1 to the rs17293632 (C) allele, but not to the sequence with the (T) allele (31 AP-1 reads for the C allele versus 1 read

for the T allele). Their findings are directionally similar to our allele-specific TaqMan experiments with AP-1 (Figure 5B). However, their study did not provide further experimental evidence with respect to rs17293632 function. Although rs17293632 was used as an example of a likely causal SNP, the specific AP-1 protein(s) investigated in their ChIP-Seq experiment was not reported.

Further evidence for rs17293632C>T as a causal SNP at this locus is provided by eQTL data for *SMAD3* expression in both whole blood and vascular tissue. We elected to focus solely on *SMAD3* for eQTL analysis because rs17293632 is located within *SMAD3* and no other genes in the region had plausible links to atherosclerosis. Although we demonstrated that rs17293632 is a *cis*-eQTL for *SMAD3* expression, we cannot rule out trans-eQTL effects. There are several genes in the vicinity of *SMAD3*/rs17293632 that may also be regulated by rs17293632, and future studies with this regard will be of interest. Although *P* values in eQTL analysis were nominally significant at $P<0.05$, they would not pass the multiple correction threshold if we probed for many genes in addition to *SMAD3*. Recent studies have shown some eQTLs are shared among tissues,[41,42] whereas others are tissue-specific.[42,43] Although we showed eQTL data in whole blood and vascular tissue, it will be important to determine whether rs17293632 is an eQTL in various other tissues because this SNP has pleiotropic effects on other diseases (Table IV in the online-only Data Supplement). In the midst of our experiments, rs17293632 was deposited in the GTEx eQTL database (http://www.gtexportal.org) as an eQTL for SMAD3 in thyroid with the same direction of effect ($P=1.9\times10^{-13}$; Figure VIII in the online-only Data Supplement).

Although these fine-mapping, bioinformatic, and experimental findings demonstrate that rs17293632C>T is a functional SNP at this locus, others may exist. Table 2 shows that this SNP is in almost perfect LD with 6 other SNPs at the *SMAD3* locus. Although the latter do not seem to significantly disrupt transcription factor binding sites or have strong regulatory chromatin modifications compared with rs17293632C>T, it is plausible that they could still affect gene regulation. As indicated in Figure 1C, the *SMAD3* locus as a whole seems to contain many gene regulatory hotspots, with several regions containing strong H3K27ac, H3K4me1, and clusters of transcription factor binding in ENCODE ChIP-Seq data.

Other reports indicate the enhancer sequence in *SMAD3* encompassing rs17293632C>T has relevance to phenotypes other than atherosclerosis. Of note, a cluster of SNPs in vicinity to rs17293632 and toward the 3′ end of the *SMAD3* gene (Figure 1C) have previously been associated with various diseases (Table IV in the online-only Data Supplement), most with immune components in their pathophysiology. These include Crohn disease, inflammatory bowel disease, asthma, and atopy.[44–48] As we have demonstrated, the *SMAD3* intron 1 enhancer is functional across numerous cell types and thus effects of rs17293632C>T are likely to apply across a wide variety of cell types expressing AP-1 family member proteins. Affecting enhancer function and subsequently *SMAD3* expression would be expected to alter TGF-β signaling with pleiotropic effects giving rise to various conditions and diseases. Similar to Crohn disease and inflammatory bowel disease, atherosclerosis has an immune/inflammatory component and macrophages and T cells play important roles in pathogenesis. It could thus be of interest to determine whether immune and inflammatory cytokines are able to activate AP-1 signaling upstream of binding to rs17292632. Because the rs17293632 SNP and 5′-TGA[G/C]TCA-3′ AP-1 recognition sequence is conserved between humans and mice, generation of mouse models with different alleles for rs17293632 or deletion of the enhancer sequence entirely may further unravel the contribution of this AP-1 regulatory mechanism. Alternatively, genome editing of the AP-1 binding site using CRISPR/Cas9 in human cell lines could be used to further clarify its role.

It is well established that SMAD3 is an important mediator of TGF-β signaling and previous studies suggest that TGF-β signaling has a protective effect in atherosclerosis.[9,10,12] However, the role of SMAD3 in the vasculature is somewhat less clear. As reported by Kobayashi et al,[49] *Smad3* null mice showed enhanced neointimal hyperplasia with decreased matrix deposition in response to femoral artery injury. *Smad3*$^{-/-}$ intima were also shown to contain more proliferating vascular SMCs and less collagen as compared with wild-type intima. In agreement with these findings here we find that in hASMCs (Figure 6), *SMAD3* knockdown increases cell viability consistent with an antiproliferative role of SMAD3. In contrast, Tsai et al[50] reported that overexpression of *Smad3* via adenoviral delivery to injured rat carotid artery or in cultured vascular SMCs increased SMC proliferation in response to TGF-β.

Irrespective of these discrepant findings, it is not immediately evident how decreased *SMAD3* expression and increased SMC proliferation in carriers of rs17293632 (T) would be associated with protection from CAD. This finding may suggest that decreased *SMAD3* expression has a beneficial effect at a specific stage of atherosclerosis or in a specific cellular/tissue context. It is known that *SMAD3* is expressed at low levels in normal artery but increases dramatically in response to injury and is elevated in atherosclerotic plaque.[11,50] Attenuation of this response in carriers of the rs17293632 (T) allele may somehow be protective. Second, although effects on vascular SMCs are plausible in terms of linking this SNP to atherosclerosis, it is possible that reduced *SMAD3* expression in another cell type in the vessel wall (endothelial, macrophage, T cell) would have an antiatherogenic effect. Our findings that expression of the *SMAD3* gene is under positive regulation by AP-1 is notable in that there are reports Jun proteins can antagonistically affect function of the SMAD3 protein,[51] suggesting dynamic interplay between these factors.

In summary, of the dozens of CAD-associated loci identified by the GWAS approach, few have been functionally characterized. Here, we finely map a newly reported GWAS locus to identify a causal SNP that represents a novel functional *cis*-acting element at the *SMAD3* locus on chromosome 15. The protective rs17293632 (T) allele disrupts a consensus AP-1 binding site, resulting in impaired AP-1 binding to the *SMAD3* intron 1 enhancer, and reduces enhancer activity that in turn correlates with lower *SMAD3* expression in blood and human plaque.

Acknowledgments

We would like to thank Dr J.G. Pickering for the kind provision of the HITC6 cells and Dr Thomas J. Kelley for provision of the SMAD3p-Luc plasmid. We also thank Anh-Thu Dang for her help with RNA isolation and reverse transcription.

Sources of Funding

This work was supported by the Heart and Stroke Foundation of Canada, T-7218 and BR-7519 (Dr McPherson).

Disclosures

None.

References

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3(11):e442. doi:10.1371/journal.pmed.0030442.
2. Schunkert H, König IR, Kathiresan S, et al; Cardiogenics; CARDIoGRAM Consortium. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011;43:333–338. doi: 10.1038/ng.784.
3. Deloukas P, Kanoni S, Willenborg C, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*. 2013;45(1):25–33. doi: 10.1038/ng.2480.
4. Nikpay M, Goel A, Won HH, et al; CARDIoGRAMplusC4D Consortium. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47:1121–1130. doi: 10.1038/ng.3396.
5. Samani NJ, Erdmann J, Hall AS, et al; WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. *N Engl J Med*. 2007;357:443–453. doi: 10.1056/NEJMoa072366.
6. Derynck R, Zhang YE. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature*. 2003;425:577–584. doi: 10.1038/nature02006.
7. Akhurst RJ, Hata A. Targeting the TGFβ signalling pathway in disease. *Nat Rev Drug Discov*. 2012;11:790–811. doi: 10.1038/nrd3810.
8. Ghosh S, Vivar J, Nelson CP, et al. Systems genetics analysis of genome-wide association study reveals novel associations between key biological processes and coronary artery disease. *Arterioscler Thromb Vasc Biol*. 2015;35:1712–1722. doi: 10.1161/ATVBAHA.115.305513.
9. Grainger DJ. Transforming growth factor beta and atherosclerosis: so far, so good for the protective cytokine hypothesis. *Arterioscler Thromb Vasc Biol*. 2004;24:399–404. doi: 10.1161/01.ATV.0000114567.76772.33.
10. Mallat Z, Tedgui A. The role of transforming growth factor beta in atherosclerosis: novel insights and future perspectives. *Curr Opin Lipidol*. 2002;13:523–529. doi: 10.1097/00041433-200210000-00008.
11. Kalinina N, Agrotis A, Antropova Y, Ilyinskaya O, Smirnov V, Tararak E, Bobik A. Smad expression in human atherosclerotic lesions: evidence for impaired TGF-β/Smad signaling in smooth muscle cells of fibrofatty lesions. *Arterioscler Thromb Vasc Biol*. 2004;24(8):1391–1396. doi: 10.1161/01.ATV.0000133605.89421.79.
12. Bobik A. Transforming growth factor-betas and vascular disorders. *Arterioscler Thromb Vasc Biol*. 2006;26:1712–1720. doi: 10.1161/01.ATV.0000225287.20034.2c.
13. van de Laar IMBH, Oldenburg RA, Pals G, et al. Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. *Nat Genet*. 2011;43(2):121–126. doi:10.1038/ng.744.
14. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21:263–265. doi: 10.1093/bioinformatics/bth457.
15. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet*. 2014;15:272–286. doi: 10.1038/nrg3682.
16. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007;39:311–318. doi: 10.1038/ng1966.
17. Heintzman ND, Hon GC, Hawkins RD, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009;459:108–112. doi: 10.1038/nature07829.
18. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012;22:1790–1797. doi: 10.1101/gr.137323.112.
19. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, Bernstein BE. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*. 2011;473:43–49. doi: 10.1038/nature09906.
20. Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57–74. doi: 10.1038/nature11247.
21. Lee JY, Elmer HL, Ross KR, Kelley TJ. Isoprenoid-mediated control of SMAD3 expression in a cultured model of cystic fibrosis epithelial cells. *Am J Respir Cell Mol Biol*. 2004;31:234–240. doi: 10.1165/rcmb.2003-0447OC.
22. Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet*. 2011;12:7–18. doi: 10.1038/nrg2905.
23. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature*. 2009;457(7231):854–859. doi: 10.1038/nature07730.
24. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet*. 2010;6:e1000888. doi: 10.1371/journal.pgen.1000888.
25. Musunuru K, Strong A, Frank-Kamenetsky M, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*. 2010;466:714–719. doi: 10.1038/nature09266.
26. Busque L, Belisle C, Provost S, Giroux M, Perreault C. Differential expression of SMAD3 transcripts is not regulated by cis-acting genetic elements but has a gender specificity. *Genes Immun*. 2009;10:192–196. doi: 10.1038/gene.2008.101.
27. Perisic L, Aldi S, Sun Y, Folkersen L, Razuvaev A, Roy J, Lengquist M, Åkesson S, Wheelock CE, Maegdefessel L, Gabrielsen A, Odeberg J, Hansson GK, Paulsson-Berne G, Hedin U. Gene expression signatures, pathways and networks in carotid atherosclerosis. *J Intern Med*. 2016;279:293–308. doi: 10.1111/joim.12448.
28. Perisic L, Hedin E, Razuvaev A, Lengquist M, Osterholm C, Folkersen L, Gillgren P, Paulsson-Berne G, Ponten F, Odeberg J, Hedin U. Profiling of atherosclerotic lesions by gene and tissue microarrays reveals PCSK6 as a novel protease in unstable carotid atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2013;33:2432–2443. doi: 10.1161/ATVBAHA.113.301743.
29. Folkersen L, Persson J, Ekstrand J, Agardh HE, Hansson GK, Gabrielsen A, Hedin U, Paulsson-Berne G. Prediction of ischemic events on the basis of transcriptomic and genomic profiling in patients undergoing carotid endarterectomy. *Mol Med*. 2012;18:669–675. doi: 10.2119/molmed.2011.00479.
30. Razuvaev A, Ekstrand J, Folkersen L, Agardh H, Markus D, Swedenborg J, Hansson GK, Gabrielsen A, Paulsson-Berne G, Roy J, Hedin U. Correlations between clinical variables and gene-expression profiles in carotid plaque instability. *Eur J Vasc Endovasc Surg*. 2011;42:722–730. doi: 10.1016/j.ejvs.2011.05.023.
31. Jackson V, Petrini J, Caidahl K, Eriksson MJ, Liska J, Eriksson P, Franco-Cereceda A. Corrigendum to "Bicuspid aortic valve leaflet morphology in relation to aortic root morphology: a study of 300 patients undergoing open-heart surgery." *Eur J Cardio-thoracic Surg*. 2012;41(2):471. doi:10.1093/ejcts/ezr168.
32. Bennett BL, Sasaki DT, Murray BW, O'Leary EC, Sakata ST, Xu W, Leisten JC, Motiwala A, Pierce S, Satoh Y, Bhagwat SS, Manning AM, Anderson DW. SP600125, an anthrapyrazolone inhibitor of Jun N-terminal kinase. *Proc Natl Acad Sci U S A*. 2001;98:13681–13686. doi: 10.1073/pnas.251194298.
33. Jang CW, Chen CH, Chen CC, Chen JY, Su YH, Chen RH. TGF-beta induces apoptosis through Smad-mediated expression of DAP-kinase. *Nat Cell Biol*. 2002;4:51–58. doi: 10.1038/ncb731.
34. Wildey GM, Patil S, Howe PH. Smad3 potentiates transforming growth factor beta (TGFbeta)-induced apoptosis and expression of the BH3-only protein Bim in WEHI 231 B lymphocytes. *J Biol Chem*. 2003;278:18069–18077. doi: 10.1074/jbc.M211958200.
35. Oliver FJ, de la Rubia G, Rolli V, Ruiz-Ruiz MC, de Murcia G, Murcia JM. Importance of poly(ADP-ribose) polymerase and its cleavage in apoptosis. Lesson from an uncleavable mutant. *J Biol Chem*. 1998;273:33533–33539. doi: 10.1074/jbc.273.50.33533.

36. Karin M, Liu Zg, Zandi E. AP-1 function and regulation. *Curr Opin Cell Biol*. 1997;9:240–246. doi: 10.1016/S0955-0674(97)80068-3.
37. Shaulian E, Karin M. AP-1 as a regulator of cell life and death. *Nat Cell Biol*. 2002;4:E131–E136. doi: 10.1038/ncb0502-e131.
38. Hess J, Angel P, Schorpp-Kistner M. AP-1 subunits: quarrel and harmony among siblings. *J Cell Sci*. 2004;117(pt 25):5965–5973. doi: 10.1242/jcs.01589.
39. Mechta-Grigoriou F, Gerald D, Yaniv M. The mammalian Jun proteins: redundancy and specificity. *Oncogene*. 2001;20:2378–2389. doi: 10.1038/sj.onc.1204381.
40. Farh KK-H, Marson A, Zhu J, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. 2015;518(7539):337–343. doi: 10.1038/nature13835.
41. Bullaughey K, Chavarria CI, Coop G, Gilad Y. Expression quantitative trait loci detected in cell lines are often present in primary tissues. *Hum Mol Genet*. 2009;18:4296–4303. doi: 10.1093/hmg/ddp382.
42. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (80-)*. 2015;348(6235):648–660. doi: 10.1126/science.1262110.
43. Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H, Ingle C, Beazley C, Gutierrez Arcelus M, Sekowska M, Gagnebin M, Nisbett J, Deloukas P, Dermitzakis ET, Antonarakis SE. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science*. 2009;325:1246–1250. doi: 10.1126/science.1174148.
44. Jostins L, Ripke S, Weersma RK, et al; International IBD Genetics Consortium (IIBDGC). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*. 2012;491:119–124. doi: 10.1038/nature11582.
45. Franke A, McGovern DP, Barrett JC, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet*. 2010;42:1118–1125. doi: 10.1038/ng.717.
46. Hinds DA, McMahon G, Kiefer AK, Do CB, Eriksson N, Evans DM, St Pourcain B, Ring SM, Mountain JL, Francke U, Davey-Smith G, Timpson NJ, Tung JY. A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. *Nat Genet*. 2013;45:907–911. doi: 10.1038/ng.2686.
47. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, von Mutius E, Farrall M, Lathrop M, Cookson WO; GABRIEL Consortium. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med*. 2010;363:1211–1221. doi: 10.1056/NEJMoa0906312.
48. Ferreira MA, Matheson MC, Tang CS, et al; Australian Asthma Genetics Consortium Collaborators. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. *J Allergy Clin Immunol*. 2014;133:1564–1571. doi: 10.1016/j.jaci.2013.10.030.
49. Kobayashi K, Yokote K, Fujimoto M, Yamashita K, Sakamoto A, Kitahara M, Kawamura H, Maezawa Y, Asaumi S, Tokuhisa T, Mori S, Saito Y. Targeted disruption of TGF-beta-Smad3 signaling leads to enhanced neointimal hyperplasia with diminished matrix deposition in response to vascular injury. *Circ Res*. 2005;96:904–912. doi: 10.1161/01.RES.0000163980.55495.44.
50. Tsai S, Hollenbeck ST, Ryer EJ, Edlin R, Yamanouchi D, Kundi R, Wang C, Liu B, Kent KC. TGF-beta through Smad3 signaling stimulates vascular smooth muscle cell proliferation and neointimal formation. *Am J Physiol Heart Circ Physiol*. 2009;297:H540–H549. doi: 10.1152/ajpheart.91478.2007.
51. Verrecchia F, Vindevoghel L, Lechleider RJ, Uitto J, Roberts AB, Mauviel A. Smad3/AP-1 interactions control transcriptional responses to TGF-beta in a promoter-specific manner. *Oncogene*. 2001;20:3332–3340. doi: 10.1038/sj.onc.1204448.

Significance

The majority of significant and replicated single nucleotide polymorphisms associated with complex diseases, such as coronary artery disease, are located in noncoding regions of the genome. However, the functional characterization of these noncoding variants is essential for further understanding the complex biology of atherosclerosis. In this study, we have used genetic and epigenetic fine mapping procedures to identify a novel functional coronary artery disease–associated intronic single nucleotide polymorphism in the *SMAD3* gene. This single nucleotide polymorphism disrupts binding of the activator protein-1 transcription factor and impairs gene enhancer activity for the human *SMAD3* promoter in primary human arterial smooth muscle cells and lowers SMAD3 levels in vivo in human whole blood and plaque tissue. Our data suggest this activator protein-1 regulation of SMAD3 is a novel and relevant mechanism in the pathogenesis of coronary artery disease. Because of the importance of transforming growth factor-β signaling in atherosclerosis, altered SMAD3 levels are expected to have important effects in the vasculature.

Arteriosclerosis, Thrombosis, and Vascular Biology

JOURNAL OF THE AMERICAN HEART ASSOCIATION

Functional Analysis of a Novel Genome-Wide Association Study Signal in *SMAD3* That Confers Protection From Coronary Artery Disease

Adam W. Turner, Amy Martinuk, Anada Silva, Paulina Lau, Majid Nikpay, Per Eriksson, Lasse Folkersen, Ljubica Perisic, Ulf Hedin, Sebastien Soubeyrand and Ruth McPherson

Arterioscler Thromb Vasc Biol. 2016;36:972-983; originally published online March 10, 2016;
doi: 10.1161/ATVBAHA.116.307294
Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1079-5642. Online ISSN: 1524-4636

Materials and Methods

Cell Culture

Primary human arterial smooth muscle cells (hASMCs) were obtained from two sources. Lot HITC6 was provided by the laboratory of Dr. JG Pickering. These cells were derived from primary cultures of smooth muscle cells from the human thoracic artery[1]. They were homozygous for the CC genotype at rs17293632 and all experiments were performed between passages 28-33. The second lot of hASMCs, 0000287836, was purchased from Lonza (catalog #CC-2571) and were heterozygous for rs17293632. These cells were derived from the ascending aorta and experiments were performed between passages 5 and 9. Both lots of hASMCs were grown in SmGM-2 Smooth Muscle Cell Basal Medium (Lonza) supplemented with 5% FBS, insulin, hFGF-B, GA-1000, and hEGF. HeLa cells were purchased from ATCC and grown in high glucose DMEM (Gibco) supplemented with 10% FBS, Penicillin-Streptomycin, and L-glutamine. HepG2 cells were also purchased from ATCC and grown in low glucose DMEM (Gibco) supplemented with 10% FBS, Penicillin-Streptomycin, and L-glutamine.

Generation of *SMAD3* Reporter Constructs

The putative enhancer sequence in intron 1 of *SMAD3* (Chr15:67,441,598-67,443,643) was first PCR amplified with primers: forward 5'-GATTGAGCTCCCTGTTCAGCATTTTGAGTTTC-3' (SacI restriction site); reverse 5'-GATTGCTAGCGAGCTATTGGAGACTGTGAGA-3' (NheI restriction site). PCR was performed on a previously genotyped genomic DNA template that had a CC genotype at rs17293632 and homozygous for all other SNPs in the amplicon. This PCR product, as well as empty pGL3-Promoter vector (Promega), were both double-digested with SacI and NheI and subsequently PCR or gel purified before ligation. Site-directed mutagenesis was used to change the allele at rs17293632 from C to T and performed using the Q5 site-directed mutagenesis kit (New England BioLabs).

To test for enhancer activity in the context of the human *SMAD3* promoter, we obtained the SMAD3p-Luc plasmid from the laboratory of Dr. Thomas Kelley[2]. This plasmid contains the human *SMAD3* promoter (fragment corresponding to -1879 to +13 of the ATG start site of the *SMAD3* gene) upstream of the firefly luciferase gene in the pGL3-Basic vector. From the pGL3-Promoter-rs17293632-C and pGL3-Promoter-rs17293632-T plasmids we PCR amplified a 347 bp segment using the following primers: forward 5'-GATTGTCGACCTGAGATGGTTGTAAATGTCCC-3' (SalI restriction site); reverse 5'- GATTGTCGACAACTGGCGGCCTTGTTCTAT-3' (SalI restriction site), PCR amplification was conducted using Q5 Hot Start High-Fidelity DNA Polymerase (New England BioLabs). These PCR products plus SMAD3p-Luc, were digested with SalI and subsequently PCR or gel purified before ligation. Correct sequences and insert orientations of all vectors were verified by Sanger sequencing.

Transfection and Luciferase Assays

For transient transfection studies in hASMCs, HITC6 cells were seeded in 6 well dishes and grown to ~85% confluence at the time of transfection. Each well was transfected with 2.5 µg of firefly luciferase construct and 2% pRL-SV40 *Renilla* (pGL3-Promoter experiments) or 2% pRL-TK *Renilla* (SMAD3p-Luc experiments). hASMCs were harvested 48 hours after transfection by washing once with 1X PBS and lysing with 500 µL of 1X Passive Lysis Buffer (Promega). For transient transfection studies with HeLa

and HepG2 cells, cells were seeded in 12 well dishes and grown to 70% confluence at the time of transfection. Each well was transfected with 500 ng of firefly luciferase construct and 2% pRL-TK *Renilla*. The HeLa and HepG2 cells were harvested 24 hours after transfection with 250 µL of 1X Passive Lysis Buffer per well. Transfections for all cell types were conducted using the Lipofectamine 3000 Reagent (Life Technologies). Lysates were centrifuged for 1 minute at 13,000 rpm at 4°C and 20 µL of supernatant were used for the luciferase assay. Luciferase readings for the hASMCs were measured with the Lumat LB 9507 luminometer (Berthold Technologies) and readings for the HeLa and HepG2 cells were measured with the GloMax luminometer (Promega).

Drug Treatments

To determine the effect of SP600125 on *SMAD3* enhancer activity, HITC6 hASMCs were seeded in 6 well dishes and grown to approximately 85% confluence. SP600125 was purchased from Cell Signaling Technology and added to cells for 1 h at a final concentration of 50 µM prior to transfection with the *SMAD3* pGL3-Promoter constructs (2.5 µg reporter per well). The HITC6 cells were harvested 24 h after transfection with 500 µL of 1X Passive Lysis Buffer per well. The final concentration of DMSO in both the control and SP600125 treatments was 0.1%.

PMA Treatments and c-Jun Knockdown

For PMA experiments both HITC6 and lot 287836 cells were seeded in 6 well plates and grown to approximately 85% confluence. At this point, media was changed for 16 hours to SmGM-2 Smooth Muscle Cell Basal Medium (Lonza) supplemented with 0.5% FBS, insulin, hFGF-B, GA-1000, and hEGF. Both lots of cells were treated with a final concentration of 50 nM PMA for the following time points: 0 h, 1 h, 2 h, and 4 h before being harvested for RNA.

For c-Jun siRNA knockdown, HITC6 cells were seeded in 6 well plates and grown to ~80% confluence. Cells were transfected with a final concentration of either 50 nM nontarget siRNA or 50 nM *JUN* siRNA (s7658, Life Technologies). 48 h after transfection, cells were harvested for RNA.

RNA for all samples was isolated using the High Pure RNA Isolation Kit (Roche). 500 ng of RNA was reverse-transcribed using the Transcriptor First Strand cDNA Synthesis Kit (Roche) using a combination of oligo(dT) and random hexamer primers. Quantitative PCR of cDNA samples was conducted using the LightCycler 480 II machine (Roche). Human *β-actin* was used as the reference gene for qPCR experiments. The following primer sequences were used for qPCR experiments: *SMAD3*: forward 5'-GCCTTCTGGTGCTCCATCTC-3' reverse 5'-AATAGCGCTGTCACTGAGGCA-3'; *JUN*: forward 5'-AGGTTCAGGGTCATGCTCTGTT-3' reverse 5'-ACGTGAAGTGACGGACTGTTCT-3'; *β-actin*: forward 5'-TCCCTGGAGAAGAGCTACGA-3' reverse 5'-ATCTGCTGGAAGGTGGACAG-3'.

Protein Expression

For all protein experiments, either HITC6 or lot 287836 cells were grown in 6 well plates. Protein samples were harvested on ice using RIPA buffer containing fresh Roche cOmplete Protease Inhibitor cocktail and Roche PhosStop Phosphatase Inhibitor cocktail. Protein concentration was determined using the Pierce BCA Protein Assay Kit. 35 µg of each denatured sample was loaded onto a 4-15% gradient SDS-PAGE gel (Bio-Rad) and run under reducing conditions. Samples were transferred to nitrocellulose

membranes (Bio-Rad) for 1 h at 100V and subsequently blocked for 1 h at room temperature using Odyssey blocking buffer (SMAD3, c-Jun, and PARP) or 5% BSA (phospho c-Jun). To confirm SP600125 reduced phosphorylation of c-Jun, the membrane was incubated with rabbit Phospho-c-Jun (Ser63) II Antibody #9261 (Cell Signaling Technology, 1:1000 dilution). Total c-Jun was probed by stripping the membrane and re-probing with rabbit c-Jun (H-79) antibody (Santa Cruz Biotechnology, sc-1694). To confirm SMAD3 knockdown in the cell proliferation assays, membranes were incubated with rabbit SMAD3 Antibody #9513 (Cell Signaling Technology, 1:1000 dilution). Apoptosis was assessed by probing with rabbit PARP Antibody #9542 (Cell Signaling Technology, 1:1000 dilution), that detects the normal 116 kDa form of PARP as well as the cleaved 89 kDa protein. For loading controls, all membranes were probed with mouse beta Tubulin Antibody D66 (GeneTex, 1:2000 dilution). IRDye secondary antibodies (LI-COR) were used at a concentration of 1:15000 and membranes were visualized with the LI-COR Odyssey imaging system.

Human Samples for *SMAD3* Whole Blood Analysis

Healthy normal weight individuals of European ancestry[3] were selected for eQTL studies on whole blood. Genomic DNA was genotyped using Affymetrix 6.0 arrays and SNP data imputed using the Phase 1 v3 release from the 1000 Genomes Project. The minor allele frequency for rs17293632 (T) is 0.21 (Table 1); 41 subjects in this cohort were homozygous for the (T) allele at rs17293632. Many more subjects were heterozygous or homozygous (C) carriers and 50 of each were selected with careful matching for age and sex to the rs17293632 TT group (Supplementary Table I). Fasting blood samples were collected in PAXgene Blood RNA tubes (PreAnalytiX/QIAGEN) and stored at -80°C prior to analysis.

Quantitative Real-Time PCR for SMAD3 Whole Blood Analysis

Whole blood RNA was extracted using the PAXgene Blood RNA kit (PreAnalytiX/QIAGEN). cDNA was synthesized using the Transcriptor First Strand cDNA Synthesis kit (Roche) and a combination of anchored oligo(dT) and random hexamer primers, with 500 ng of patient whole blood RNA as the template. Quantitative PCR for *SMAD3* mRNA was performed with SYBR Green I Master (Roche) and run on the LightCycler 480 II machine (Roche). Peptidylprolyl Isomerase A (*PPIA*) was used as the reference gene for all quantitative PCR experiments. The following primer sequences were used for quantitative PCR analysis: *SMAD3*: forward 5'-GCCTTCTGGTGCTCCATCTC-3' reverse 5'-AATAGCGCTGTCACTGAGGCA-3'; *PPIA*: forward 5'-ACCGTGTTCTTCGACATTGC-3' reverse 5'-TTCTGTGAAAGCAGGAACCC-3'

***SMAD3* eQTL Analysis in Primary Tissue**

SMAD3 expression in human carotid plaque tissue was investigated using samples from the BiKE (Biobank of Karolinska Endarterectomies) study. The BiKE study at the Karolinska Institute, Stockholm, Sweden comprises a repository of atherosclerotic plaque and plasma samples obtained from patients that underwent carotid endarterectomy, along with clinical parameters, genotype data, RNA expression and proteomics data. Details of the study cohort have been published previously[4–7]. High-density genotyping in BiKE samples was performed using Illumina610w-QuadBead Array and matched with the corresponding plaque samples (n=125). Since rs17293632 was not present on the chip, rs16950687 (D'

0.949, r^2 0.859 with rs17293632; D' 1, r^2 0.953 with rs56062135) was selected as a proxy using the SNaP software from the Broad Institute.

SMAD3 expression was also probed *in vivo* in non-CAD patients from various tissues collected in the ASAP (Advanced Study of Aortic Pathology) study. Patients from the ASAP study had aortic valve and ascending aortic disease and all underwent elective open heart surgery[8]. None of these patients had significant CAD by coronary angiography. Tissues collected included mammary artery intima-media, liver, ascending aorta intima-media, ascending aorta adventitia, and heart. SNPs were genotyped on the Illumina610w-QuadBead Array and rs17293632 imputed using the Mach2.0 algorithm and 1000 Genomes data as reference (imputation quality $r^2 = 0.718$).

Chromatin Immunoprecipitation

Primary hASMCs were seeded in 150 mm dishes and grown to confluence before harvesting. Cells were fixed with 1% formaldehyde before being quenched with glycine. For each round of ChIP, cells from two 150 mm plates (~1 million cells per dish) were combined and lysed with 500 µL of nuclear lysis buffer. Cross-linked chromatin was sheared into fragments of approximately 500 bp using the BioRuptor sonicator (Diagenode) and clarified using centrifugation. PureProteome Protein A/G Magnetic Beads (Millipore) were preblocked with both salmon sperm DNA and BSA for 30 minutes at 4°C. Before the immunoprecipitation, the 500 µL of total chromatin was incubated with 10 µL of resuspended pre-blocked beads to further reduce background. Each chromatin immunoprecipitation used 200 µg of hASMC chromatin, 10 µL of resuspended beads, and 10 µg of either c-Fos (Santa Cruz Biotechnology, K-25: sc-253), c-Jun (Santa Cruz Biotechnology, H-79: sc-1694), JunB (Santa Cruz Biotechnology, 210: sc-73), JunD (Santa Cruz Biotechnology, 329: sc-74), P300 (Santa Cruz Biotechnology, C-20: sc-585), H3K27ac (Abcam, ChIP Grade ab4729), or nonspecific Rabbit IgG. Protein-antibody-bead complexes were incubated overnight at 4°C with rotation. Complexes were washed four times with wash buffer, once with high salt wash buffer, and eluted from the beads with elution buffer containing 1% SDS and 100 mM sodium deoxycholate and heating for 15 minutes at 65°C. Cross-links were reversed by addition of NaCl to a final concentration of 200 mM along with RNase A and shaking at 65°C for 5 hours. Eluted DNA was purified using the GeneJET PCR Purification Kit (Thermo Scientific) and used for quantitative PCR. Quantitative PCR was performed using SYBR Green I Master (Roche) and the LightCycler 480 II machine (Roche). The sequences of primers specific for rs17293632 were FWD: 5'-CTCCGCGTGAATGTCACTG-3' and REV: 5'- GAGAGGTGAAGAGGGCAAAT-3'. Sequences of negative control primers, amplifying intronic areas of the *CPT1B* gene, are listed in Supplementary Table II. Melting curve analysis was conducted for each primer set for every run conducted.

Allele-Specific ChIP-qPCR

A TaqMan SNP Genotyping Assay for the rs17293632 SNP was purchased from Life Technologies (C__33991343_10). Lot 287836 hASMCs, confirmed to be heterozygous at rs17293632, were used for all allele-specific ChIP qPCR experiments. Quantitative PCR was first performed on the immunoprecipitated DNA as described above, and products purified using the GeneJET PCR Purification Kit (Thermo Scientific). TaqMan SNP genotyping assays for rs17293632 were then performed on PCR products to generate the ratio of the (C) to (T) allele for DNA that binds AP-1. For each allele-specific qPCR run, a standard curve was generated consisting of previously genotyped homozygous CC and homozygous TT

genomic DNA samples mixed in the following ratios: 9:1, 8:2, 7:3, 6:4, 5:5, 4:6, 3:7, 2:8, and 1:9. The Log2 ratio of FAM/VIC of these standards at cycle 55 was then used to generate the standard curve. The Log2 ratio of FAM/VIC of the immunoprecipitated DNA was then fitted to the standard curve to calculate the ratio of (C) to (T) alleles. Both sonicated input PCR product, as well as unsonicated lot 287836 genomic DNA, which have 1:1 C:T ratios, were included as controls for each run.

Cell Proliferation Assays

Cell proliferation assays were performed using the same protocol for HITC6 and lot 287836 batches of hASMCs. These were first seeded in 24 well plates and grown in SmGM-2 Smooth Muscle Cell Basal Medium (Lonza) supplemented with 5% FBS, insulin, hFGF-B, GA-1000, and hEGF. siRNA transfection was performed when cells reached approximately 45% confluency. Cells were transfected with a final concentration of either 20 nM nontarget siRNA or 20 nM *SMAD3* siRNA. *SMAD3* siRNA oligonucleotides (s8401 and s8402) were purchased from Life Technologies. siRNA transfection was conducted using the Lipofectamine RNAiMAX reagent (Life Technologies) and 6 h after transfection, media was changed and replaced with normal growth medium. After 72 h of knockdown, cell proliferation was measured using the MTT assay (Molecular Probes, Life Technologies). The MTT assay detects living cells, and is based on reduction of the tetrazolium dye MTT to its insoluble form, formazan, that has a purple colour that can be quantified at 595 nm. After MTT incubation at 37°C, DMSO was added as the solubilizing reagent.

Conditional and Joint Association Analysis

To determine if the observed association between *SMAD3* and CAD is mediated by rs56062135 or other SNPs, we used conditional and joint association analysis implemented in GCTA software[9]. This is a two step procedure that involves conditional analysis to identify independent SNPs and secondly, fitting all selected SNPs jointly in a model and excluding SNPs with p values (p-values from joint analysis) that are greater than the cut-off p value, using the software default $p< 5x\ 10^{-8}$ (GWAS significant level) as our threshold. GCTA method requires summary-level statistics from a meta-analysis of genome-wide association studies (GWAS) and estimated linkage disequilibrium (LD) from a reference sample with individual-level genotype data. We used the SNP summary statistics from the recently published 1000 Genomes–based GWAS meta-analysis of CAD[10] as our meta-analysis file (CARDIoGRAM Meta-analysis). Furthermore, the reference sample must be large enough so that the LD correlations are estimated with little error[9,11], and it can be one of the participating studies of the meta-analysis[9]; therefore, we used the 1000 Genomes post-imputed genotype data from our cohorts that are included in the CARDIoGRAM Meta-analysis and excluded variants with IMPUTE2 info < 0.4, in our final dataset, we have 13,367 subjects with available genotypes at 15M variants that we used as our reference sample for conditional and joint association analysis.

Statistical Analysis

Data displayed represents the mean ± SD for experiments conducted in triplicate, and is representative of 3 independent experiments. Significant differences were assessed by either one-way ANOVA followed by Bonferroni's post-hoc test or unpaired, two-tailed Student's T-test, where appropriate. Figures were generated and analysis performed using the GraphPad Prism software.

References

1. Li S, Fan Y-S, Chow LH, Van Den Diepstraten C, van der Veer E, Sims SM, Pickering JG. Innate diversity of adult human arterial smooth muscle cells: cloning of distinct subtypes from the internal thoracic artery. *Circ Res*. 2001;89(6):517-525. doi:10.1161/hh1801.097165.

2. Lee JY, Elmer HL, Ross KR, Kelley TJ. Isoprenoid-mediated control of SMAD3 expression in a cultured model of cystic fibrosis epithelial cells. *Am J Respir Cell Mol Biol*. 2004;31:234-240. doi:10.1165/rcmb.2003-0447OC.

3. Cole CB, Nikpay M, Lau P, Stewart AFR, Davies RW, Wells GA, Dent R, McPherson R. Adiposity significantly modifies genetic risk for dyslipidemia. *J Lipid Res*. 2014;55(11):2416-2422. doi:10.1194/jlr.P052522.

4. Perisic L, Aldi S, Sun Y, Folkersen L, Razuvaev A, Roy J, Lengquist M, Akesson S, Wheelock CE, Maegdefessel L, Gabrielsen A, Odeberg J, Hansson GK, Paulsson-Berne G, Hedin U. Gene expression signatures, pathways and networks in carotid atherosclerosis. *J Intern Med (in Press)*. 2015.

5. Perisic L, Hedin E, Razuvaev A, Lengquist M, Osterholm C, Folkersen L, Gillgren P, Paulsson-Berne G, Ponten F, Odeberg J, Hedin U. Profiling of atherosclerotic lesions by gene and tissue microarrays reveals PCSK6 as a novel protease in unstable carotid atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2013;33(10):2432-2443. doi:10.1161/ATVBAHA.113.301743.

6. Folkersen L, Persson J, Ekstrand J, Agardh HE, Hansson GK, Gabrielsen A, Hedin U, Paulsson-Berne G. Predicted of ischemic events on the basis of transcriptomic and genomic profiling in patients undergoing carotid endarterectomy. *Mol Med*. 2012;18(4):669-775. doi:10.2119/molmed.2011.00479.

7. Razuvaev A, Ekstrand J, Folkersen L, Agardh H, Markus D, Swedenborg J, Hansson GK, Gabrielsen A, Paulsson-Berne G, Roy J, Hedin U. Correlations between clinical variables and gene-expression profiles in carotid plaque instability. *Eur J Vasc Endovasc Surg*. 2011;42(6):722-730. doi:10.1016/j.ejvs.2011.05.023.

8. Jackson V, Petrini J, Caidahl K, Eriksson MJ, Liska J, Eriksson P, Franco-Cereceda A. Corrigendum to "Bicuspid aortic valve leaflet morphology in relation to aortic root morphology: A study of 300 patients undergoing open-heart surgery." *Eur J Cardio-thoracic Surg*. 2012;41(2):471. doi:10.1093/ejcts/ezr168.

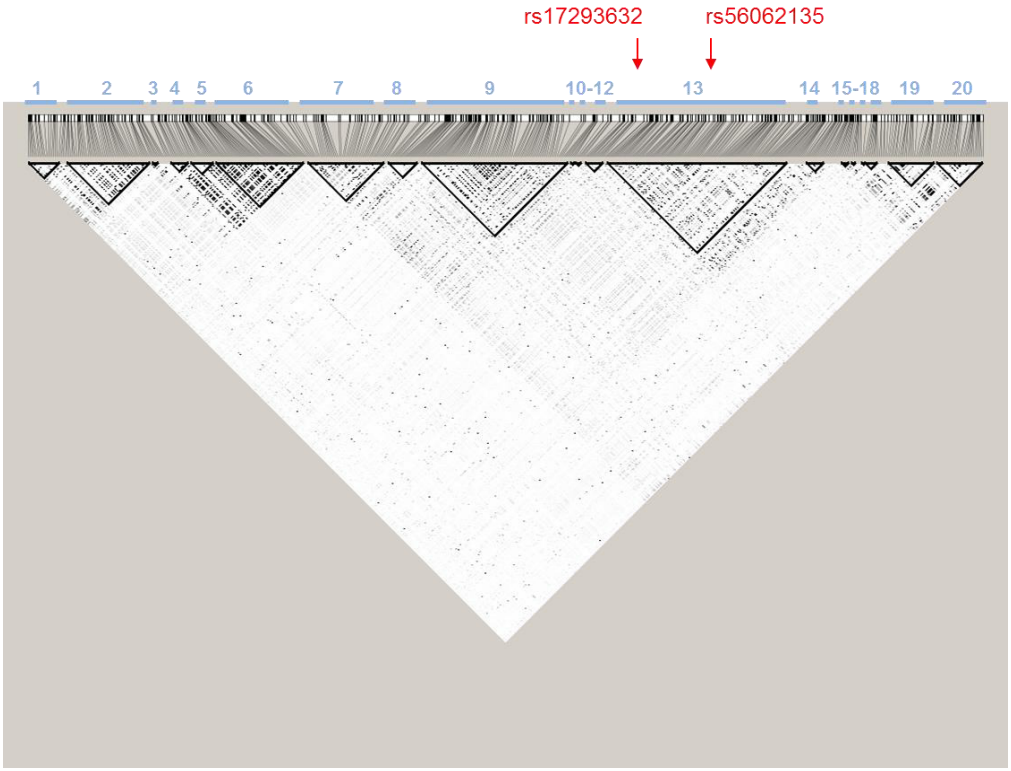
9. Yang J, Ferreira T, Morris AP, Medland SE, Madden PAF, Heath AC, Martin NG, Montgomery GW, Weedon MN, Loos RJ, Frayling TM, McCarthy MI, Hirschhorn JN, Goddard ME, Visscher PM. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet*. 2012;44(4):369-375. doi:10.1038/ng.2213.

10. Nikpay M, Goel A, Won H-H et al. A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47(10):1121-1130. doi:10.1038/ng.3396.

11. Pardo L, Bochdanovits Z, de Geus E, Hottenga JJ, Sullivan P, Posthuma D, Penninx BWJH, Boomsma D, Heutink P. Global similarity with local differences in linkage disequilibrium between the Dutch and HapMap-CEU populations. *Eur J Hum Genet*. 2009;17(6):802-810. doi:10.1038/ejhg.2008.248.

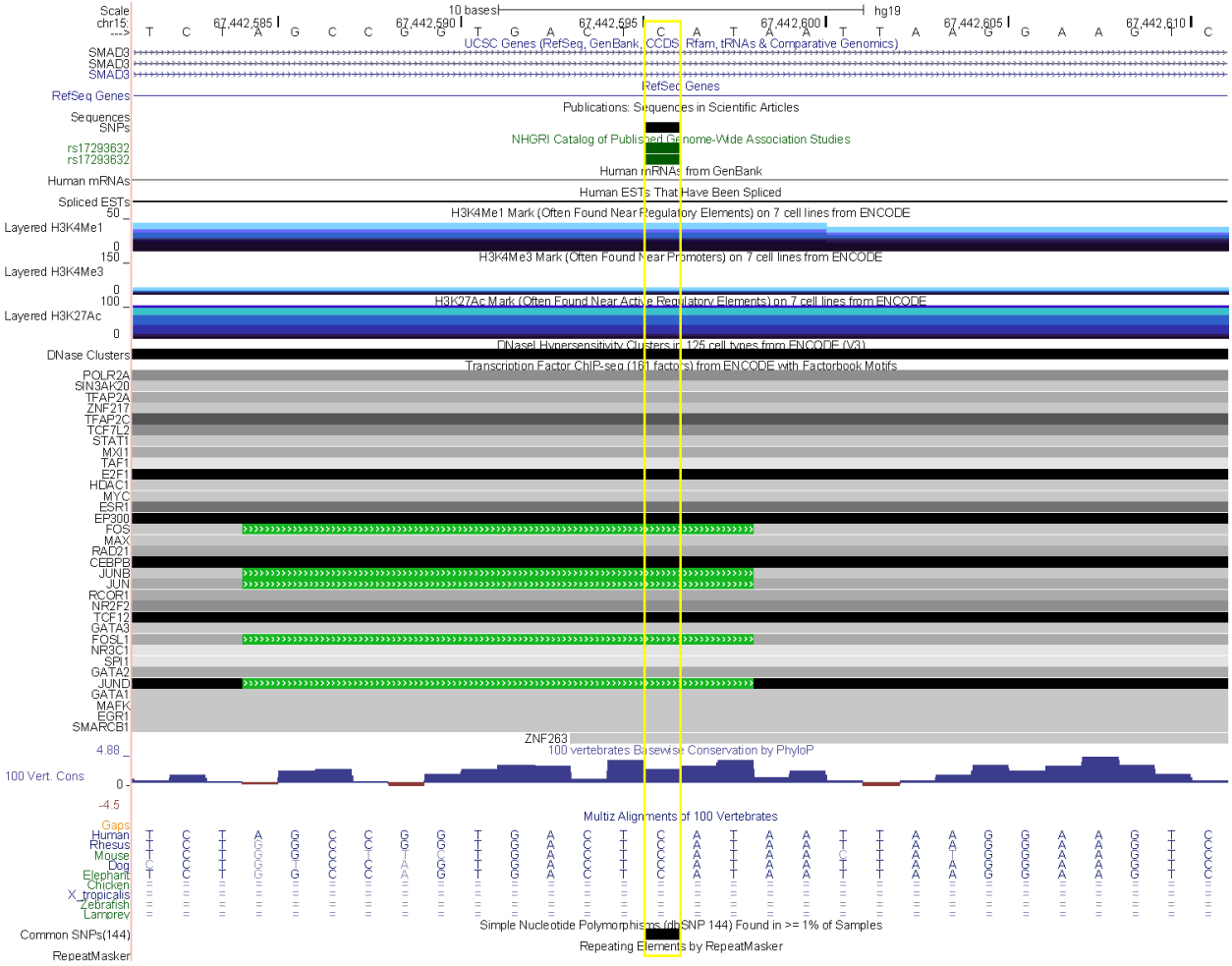
Supplement Material

Supplementary Figure I



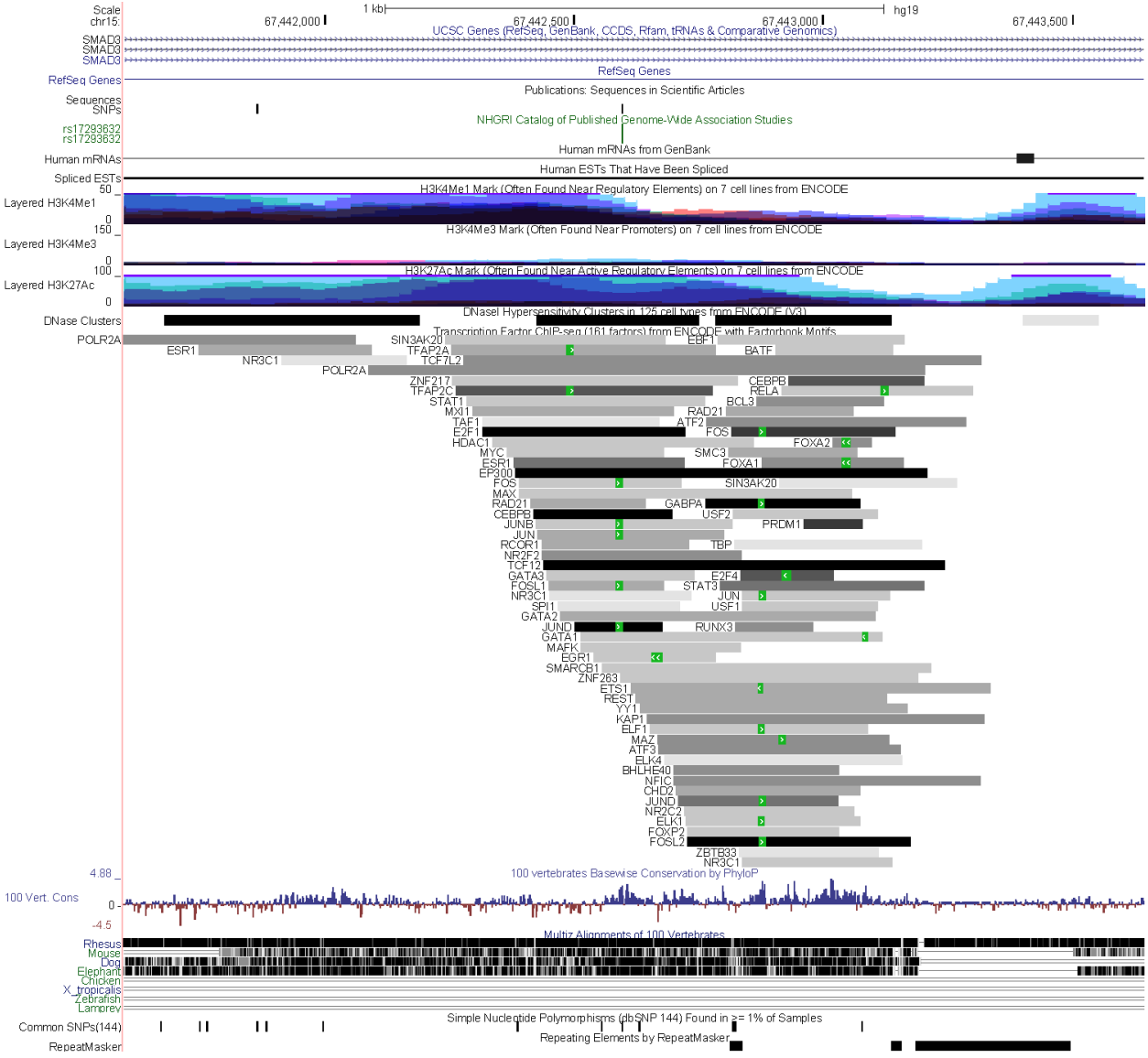
Supplementary Figure I. Visualization of the various LD blocks in the *SMAD3* gene using the Haploview software and linkage data taken from the 1000 Genomes Browser.

Supplementary Figure II



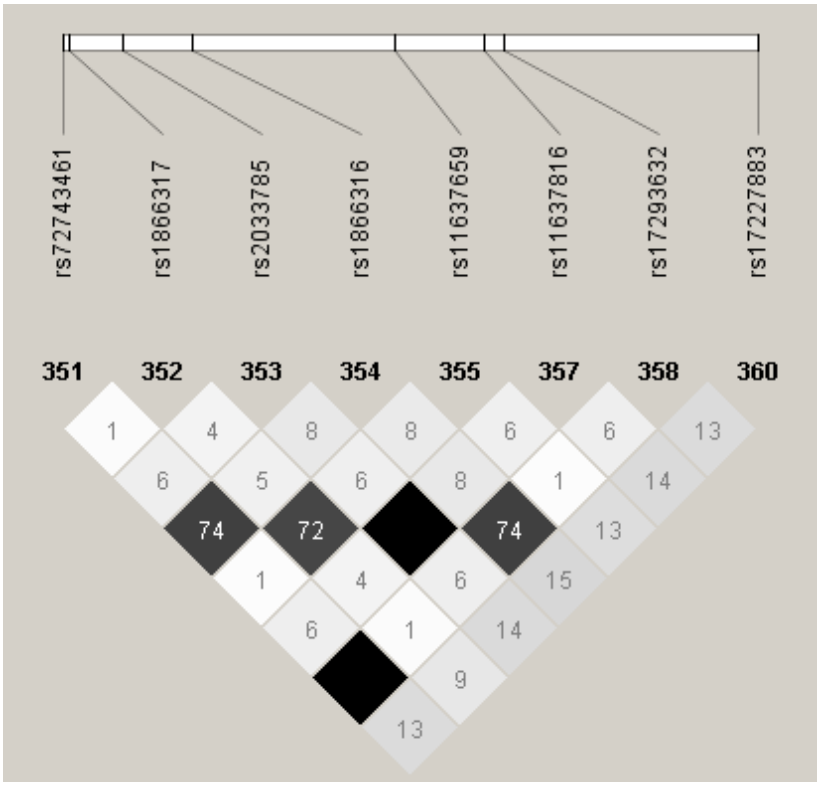
Supplementary Figure II. UCSC genome browser annotation and ENCODE project data for the regional view of the rs17293632 SNP.

Supplementary Figure III



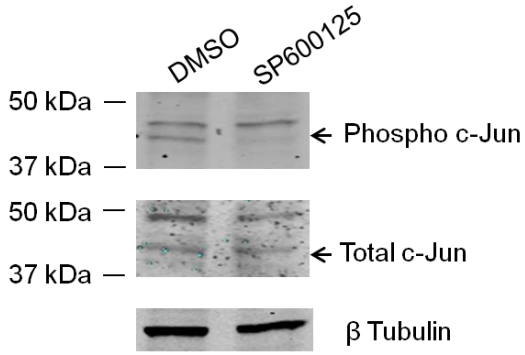
Supplementary Figure III. UCSC genome browser annotation and ENCODE project data for the entire full *SMAD3* intron 1 enhancer sequence characterized.

Supplementary Figure IV



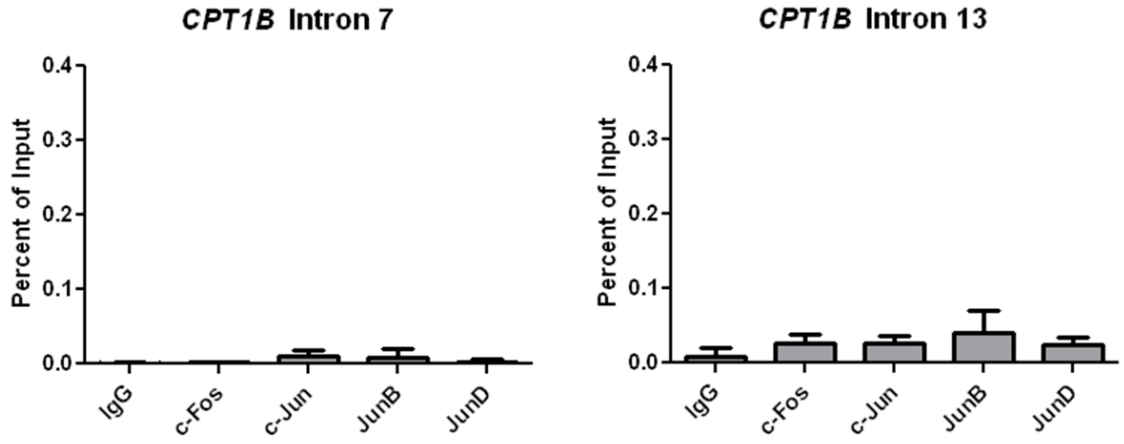
Supplementary Figure IV. LD heatmap generated using the Haploview software and linkage data taken from the 1000 Genomes Browser displaying pairwise linkage disequilibrium for SNPs located within the *SMAD3* intron 1 enhancer. Numbers within the blocks represent pairwise r^2 values multiplied by 100. White to grey to black blocks represent increasing r^2 values.

Supplementary Figure V



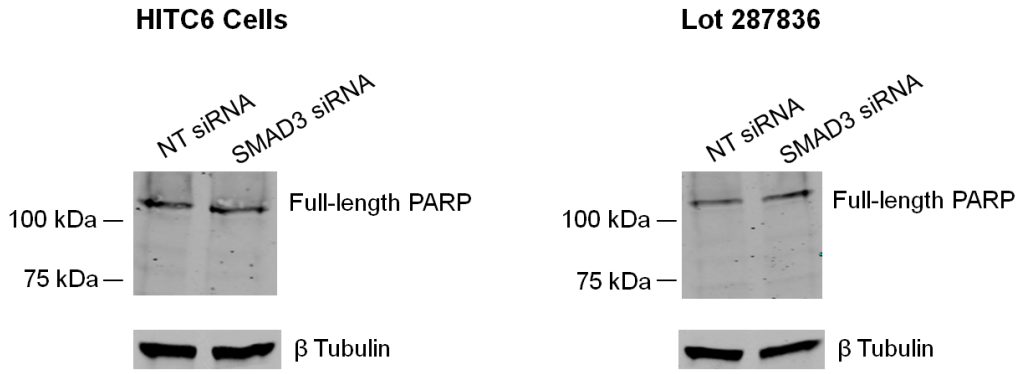
Supplementary Figure V. Reduced levels of phospho c-Jun in response to SP600125 were confirmed via Western blotting, with human β-tubulin used as the loading control. HITC6 hASMCs were treated with either 50 μM SP600125 or DMSO control for 24 hours prior to harvesting.

Supplementary Figure VI



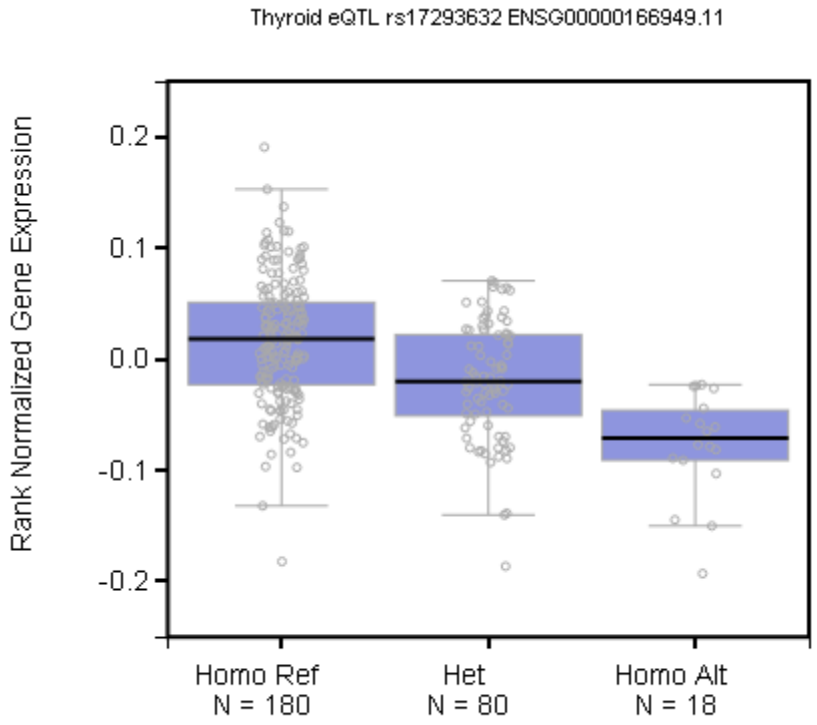
Supplementary Figure VI. Negative control DNA regions amplified for AP-1 chromatin immunoprecipitation. Crosslinked DNA-protein complexes from HITC6 cells were immunoprecipitated with the antibodies for AP-1 components c-Fos, c-Jun, JunB, and JunD. The ChIP DNA was then amplified with primers specific for two regions of the *CPT1B* gene (intron 7 and intron 13). Values are expressed as the percentage of input chromatin DNA. Values denote the mean of three independent ChIP experiments ± SD.

Supplementary Figure VII



Supplementary Figure VII. *SMAD3* siRNA knockdown does not affect cleavage of PARP in both lots of hASMCs. HITC6 and lot 287836 cells were transfected with either 20 nM nontarget siRNA or 20 nM *SMAD3* siRNA and harvested for protein 72 hours later. Lysates were probed with PARP antibody that recognizes both the full length, 116 kDa form of the protein, as well as the 89 kDa form that serves as a marker for cells undergoing apoptosis. Human β-tubulin was used as the loading control for both lots of cells.

Supplementary Figure VIII



Supplementary Figure VIII. eQTL box plot for rs17293632 and SMAD3 mRNA expression in thyroid tissue from the GTEx eQTL Browser. ENSG00000166949.11 refers to SMAD3 levels.

Supplementary Table I. Characteristics of the three groups of healthy individuals selected for rs17293632 eQTL analysis

Cohort	N Male	N Female	Mean Age (± SD)	Mean BMI (± SD)
CC	25	25	45.68 (15.78)	21.00 (1.70)
CT	25	25	43.68 (14.30)	20.20 (1.98)
TT	19	22	44.65 (15.82)	20.60 (1.80)

Supplementary Table II. Sequences of the negative control intronic DNA regions used for AP-1 chromatin immunoprecipitation

Region	Amplicon Size	Forward primer	Reverse primer
CPT1B Intron 7	123	GGAAGAGGGAGACAGAACACC	GCACAAGGTCCTACAGAGGAA
CPT1B Intron 13	101	TCAGCACTTCCCTCTTACCC	GGCCAAGACATCTTCAGGAG

Supplementary Table III. Conditional and joint association analysis results for *SMAD3* locus

		CARDIoGRAM Meta-analysis			Conditional and joint analysis		
Variant	Effect allele	Beta	SE	p-value	Beta	SE	p-value
rs56062135	C	0.07	0.01	4.52E-09	0.07	0.01	4.54E-09

Supplementary Table IV. SNPs within the 2 kb enhancer in intron 1 (chr15:67,441,598-67,443,643) of the *SMAD3* gene that are in linkage disequilibrium with published GWAS SNPs

Reported GWAS SNP (MAF)	Disease or Trait	Reference	Enhancer SNP (MAF)	D'	r^2
rs56062135 (0.21)	Coronary artery disease	Nikpay et al. *Nat Genet* 2015	rs72743461 (0.22)	0.97	0.92
			rs1866316 (0.28)	0.97	0.66
			rs17293632 (0.21)	0.97	0.94
rs17228212 (0.29)	Coronary artery disease	Samani et al. *N Engl J Med* 2007	rs17227883 (0.29)	0.98	0.96
rs12441344 (0.22)	Risk of sudden cardiac arrest in patients with coronary artery disease	Tseng et al. *Heart Rhythm* 2009	rs2033785 (0.23)	0.98	0.90
			rs11637816 (0.23)	0.98	0.90
rs17293632 (0.21)	Inflammatory bowel disease	Jostins et al. *Nature* 2012	rs72743461 (0.22)	1	1
			rs1866316 (0.28)	1	0.70
			rs17293632 (0.21)	1	1
rs17293632 (0.21)	Crohn's disease	Franke et al. *Nat Genet* 2010	rs72743461 (0.22)	1	1
			rs1866316 (0.28)	1	0.70
			rs17293632 (0.21)	1	1
rs17228058 (0.21)	Allergy susceptibility	Hinds et al. *Nat Genet 2013*	rs72743461 (0.22)	0.96	0.92
			rs1866316 (0.28)	0.96	0.65
			rs17293632 (0.21)	0.97	0.93
rs744910 (0.52)	Asthma	Moffat *N Engl J Med* 2010	rs72743461 (0.22)	-0.98	0.29
			rs2033785 (0.23)	0.99	0.27
			rs1866316 (0.28)	-0.97	0.40
			rs11637659 (0.80)	1	0.27
			rs11637816 (0.23)	0.99	0.27
			rs17293632 (0.21)	-0.98	0.28
			rs17227883 (0.29)	0.96	0.35

rs12913547 (0.24)	Central corneal thickness and keratoconus	Lu et al. *Nat Genet* 2013	rs2033785 (0.23)	0.84	0.70
			rs11637816 (0.23)	0.84	0.70
rs17294280 (0.24)	Asthma and hay fever	Ferreira et al. *J Allergy Clin Immunol* 2014	rs72743461 (0.22)	0.88	0.66
			rs1866316 (0.28)	0.74	0.46
			rs17293732 (0.21)	0.68	0.89